ERDC/TEC TR-08-2

Topographical Engineering Center

US Army Corps of Engineers®
Engineer Research and Development Center

A Survey of Terrain Modeling Technologies and Techniques

Nick Gorkavyi, Jerry Snyder, and J. David Lashlee

September 2007

ERDC/TEC TR-08-2
September 2007

A Survey of Terrain Modeling Technologies and Techniques

Nick Gorkavyi

Computational Computing Services
Haymarket, VA

Jerry Snyder

General Dynamics Armament and Technical Products
Charlotte, NC

Jon D. Lashlee

Topographical Engineering Center (TEC)
7701 Telegraph Road
Alexandria, VA 22135-3864

Final Report

Approved for public release; distribution is unlimited.

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

Abstract: Test planning, rehearsal, and distributed test events for Future Combat Systems (FCS) require rapid generation of high-fidelity synthetic environments. These environments consist of high resolution synthetic scenes of test and training ranges, which use high and low resolution digital terrain surface models, 2-D and 3-D surface objects and other geospatial data to replicate site conditions. The largest component of developing synthetic 3-D scenes is the commercially available Interferometric Synthetic Aperture Radar (IFSAR) and LIght Detection And Ranging (LIDAR) data collected from airborne platforms. These industries are seeing rapid growth in data availability, and numbers and types of sensors, but the commercially available software used to process these types of data to provide high resolution, high accuracy topographic products is limited in its ability to process and produce data quickly and accurately for FCS. This work assessed the fidelity and quality of the commercial digital surface model (DSM) and digital terrain model (DTM) from Intermap; developed algorithms based on automated feature extraction (AFE) for LIDAR data that can be applied to processing IFSAR data (to perform foliage/vegetation removal and building/structure filtering while maintaining accurate terrain profile), and assessed improvements in constructing DSM/DTM using the Computational Consulting Services (CCS)-developed methods.

Executive Summary

The Army Future Combat Systems (FCS) and other Army system-of-system programs are highly dependent on Modeling and Simulation (M&S) for Developmental and Operational Testing, experimentation, and embedded training. This requires the rapid generation and regular maintenance of high-fidelity synthetic environments, Level 5 or better Digital Terrain Elevation Data (DTED), Digital Terrain Models (DTM), and geospatial databases constructed to allow efficient transmission and exchange over network-centric architectures and low-bandwidth communications networks. A challenge to constructing and maintaining high-fidelity terrain and geospatial databases in support of the Future Force involves the collection and processing of a variety of remote sensing data. From 2004 to 2006, the Army Test and Evaluation Command (ATEC) and Topographic Engineering Center (TEC) led a multi-phase project to accomplish the following for the FCS Combined Test Organization (CTO):

- **Phase 1:** Establish a set of data collection standards and requirements for Light Detection and Ranging (LIDAR) sensors and photogrammetric imagery that would yield data, independent of sensor manufacturer or platform, for constructing DTMs and geospatial databases with consistent fidelity and attribution
- **Phase 2:** Assess the capability of autonomous algorithms to identify and characterize surface objects and features using color imagery collected from aerial platforms
- **Phase 3:** Develop and optimize automated methods and techniques to generate DTED Level 5 or better DTMs and three-dimensional (3-D) models of surface features using data from LIDAR sensors and mathematical operations to describe complex geospatial data objects and 3-D topology in highly-compact manners
- **Phase 4:** Assess the capabilities and limitations of using commercial Interferometric Synthetic Aperture Radar (IFSAR) and applying techniques developed in Phases 2 and 3 to improve the accuracy of IFSAR-derived DTMs and geospatial databases.

Prior to commencement of this project, the Government collected LIDAR data for a 100-km² area of White Sands Missile Range (WSMR) based on an anecdotal set of requirements for digital terrain construction. These LIDAR data were collected based on the vendor's capabilities and did not reflect the requirements for construction of DTM and geospatial databases

for FCS M&S. In **Phase 1**, ATEC contracted with General Dynamics Armament and Technical Products (GDATP), the Greenwich Institute for Science and Technology (GIST), and Computational Consulting Services, Inc. (CCS) to analyze and then process these LIDAR data and identify deficiencies in the data that affect the overall quality and usability for constructing high-fidelity DTMs and geospatial models. Upon successful analyses of the collected data, specific requirements were established for: (1) pre-collection activities such as number and placement of ground control points, aircraft flight pattern plans, and instrument calibration; (2) raw data collection and data pre-processing including point density, collection controls, information content, and methods for independent validation and verification; and (3) data delivery formats and database structures. These requirements are independent of the individual LIDAR system and service vendor, which provides the Government with the ability to contract for LIDAR data collection while ensuring accuracy and consistency of DTMs and geospatial models derived from that data. Included in the Requirements Document is a reference guide for specifying data collection requirements and the effect of these requirements on cost, data accuracy and quality, and terrain model resolution and fidelity. The details of Phase 1 are documented in "LIDAR and Imagery Data Collection Requirements to Support the Generation of Synthetic Scenes and Digital Terrain Models" (January 2005).

The Government also collected high resolution color imagery prior to the commencement of this project. Manually processing and analyzing imagery data and extracting objects is a very time-consuming process that often yields inconsistent results. ATEC recognized the need for an automated toolset that provided the capability to process these large datasets to identify and characterize surface features such as foliage, roads, and buildings included in 3-D geospatial databases. In **Phase 2**, the GDATP/GIST/CCS team analyzed the color imagery data for more than 29 km^2 and developed automated methods to accurately identify and delineate vegetation. The team also demonstrated the ability to extract road features from imagery by processing data for a 4.5 km^2 area, generating polylines for 292 roads totaling more than 42 km. Phase 2 included a series of tasks to refine the capability to extract of foliage against background with variability of brightness and spectral parameters and estimate the foliage height for different types of vegetation. Phase 2 results were provided in "Processing Color Imagery for White Sands Missile Range" (May 2005).

Phase 3 employed optimized techniques and methods for processing LIDAR and color imagery data to construct a high-fidelity DTM, 3-D geospatial model, and attributed geospatial database for a 10 km x 10 km area at WSMR known as the Common Operating Area for FCS Distributed Test Event (DTE) #4. The team delivered a bare-earth DTED Level-5 DTM (1-m resolution) for the entire area and a DTED Level-6 DTM (30-cm resolution) for an inset area of approximately 15 km^2. The team also delivered a LIDAR intensity image of the terrain with same resolution as DTM and vegetation models in two forms: pixel level and geometrical primitives (ellipsoids). The foliage database included 3.1 million objects that were extracted from 30-cm resolution data and 1.7 million objects extracted from 1-m resolution data. The foliage data were delivered using a newly developed method using ellipsoids and manipulating each object with ellipsoidal functions and algorithms to obtain a highly compact database without compromising accuracy or information content. Phase 3 also yielded documented techniques for tracking and estimating error sources within a DTM and other geospatial data layers, including budget of errors for each pixel. The team developed filtering techniques to compensate for point-like artifacts, such as holes and holidays, by using a bare-earth DTM having a smooth, predictable surface. The ATEC project team led the development of techniques for characterizing sensor noise and capturing that information in a geospatial database. The results of Phase 3 showed that DTMs can be rapidly constructed using automated tools with elevation accuracy of 30 to 100 cm when LIDAR and color imagery are collected in accordance with the Phase 1 Requirements Document. Phase 3 also demonstrated the capability to integrate DTMs with varying levels of detail, accuracy and resolution into a single model – a critical element of maintaining data currency.

During Phase 3, the team validated the Requirements Document generated in Phase 1 along with the quality assurance metrics using the data collected prior to the start of this project. The validation processes demonstrated that results varied greatly from vendor to vendor without a single set of requirements. The disparities between vendors and their systems adversely affected the quality and accuracy of the derived DTM and geospatial data. The foliage/vegetation models were also significantly affected by the presence of artifact data, which limited their usefulness for FCS M&S and employment for the ATEC DTE.

Phase 3 results are documented in a technical report: “Generation of Synthetic Scenes and Digital Terrain Models Using LIDAR Data at White Sands Missile Range” (February 2005). The results of Phase 3 were also published by SPIE: “High Fidelity Terrain Models and Geospatial Datasets for Use in

Distributed Test Environments" (Snyder, G., Gorkavyi, N., Lashlee, D., and Lorenzo M. April 2005).

IFSAR data has been collected by the National Geospatial Agency (NGA) for large areas of the United States through commercial contracts. TEC identified the value of using IFSAR data to develop high resolution terrain models. In **Phase 4** of this project, the team performed Verification and Validation (V&V) of commercial IFSAR-derived DTM collected by Intermap Technologies NEXTMap USA 3-D Mapping Program and acquired from NGA for an area of Yuma Proving Ground called the West Fork sub-basin of Yuma Wash. In addition, the team conducted a quality and accuracy assessment of the raw commercial IFSAR and investigated methods for improving an IFSAR-derived DTM and geospatial data models. The Phase 4 assessment included:

- **Elevation Accuracy**: an analysis of the overall elevation accuracy measurements by determining the magnitude of error between a set of ground-truth control points (Champion and Lashlee 2007) and the Digital Surface Model (DSM) and DTM provided by the vendor using Root Mean Square Error (RMSE) metrics. Using the RMSE not only analyzes the accuracy but also identifies potential shifts or biases in the data
- **Percentage of Extracted Artificial Objects and Foliage:** performed foliage/vegetation removal and building/structure filtering measured by percent of extracted artificial objects and foliage
- **Percentage of Extracted Details of Relief:** measurements for maintaining accurate terrain profile as a percentage of maintained details of relief.

The overall objectives of Phase 4 were:

1. Assess the fidelity and quality of the commercial DSM and DTM from Intermap for the West Fork sub-basin of Yuma Wash
2. Develop algorithms based on Automated Feature Extraction (AFE) for LIDAR data that can be applied to processing IFSAR data – specifically to perform foliage/vegetation removal and building/structure filtering while maintaining accurate terrain profile
3. Assess improvements in constructing DSM/DTM using the CCS-developed methods.

The team received the Government Furnished IFSAR data and control points for the West Fork sub-basin of Yuma Wash, Yuma Proving Ground,

and determined that there were relatively large biases were found when compared to Government control points. For 12 percent of control points the bias was approximately 1 m. Analysis of the DSM and DTM product from Intermap concluded:

- The DSM delivered by Intermap compared to 2,481 Government surveyed control points showed an elevation accuracy of 1.17 m RMSE and the Intermap DTM had an elevation accuracy of 1.26 m RMSE for the same control points
- The commercial DSM and DTM from Intermap derived from IFSAR data had a bias of nearly 1 m for more than 2,186 control points or nearly 88 percent of the data surveyed
- Analysis also showed that the Intermap DTM is further degraded with increase foliage density and terrain gradient or offsets
- Development and refinement of algorithms for identifying and quantifying errors and biases in the Intermap DSM and DTM were successful
- The results also showed that custom algorithms derived under this project significantly improved the accuracy of the Intermap DSM and DTM by compensating for measurement biases using control-point data.

Under Phase 4, the team development and testing of AFE methods and algorithms tailored for IFSAR and used the software to construct a DTM for 25 km^2 – referred to herein as the Enhanced DTM:

- Results show that the AFE and other terrain/surface construction methods developed under Phases 2 – 4 can significantly improve the accuracy of the IFSAR derived DTM. The team was able to successfully remove surface objects and foliage without adversely affecting spatial accuracy
- The AFE methods developed in Phase 4 will provide substantial improvements to IFSAR derived DTMs in areas of heavy vegetation
- Across the set of control points, implementation of the AFE algorithms reduced the errors in the DTM from 35 to 50 percent. For the 2,186 control points, implementation of the AFE algorithms to process the IFSAR data reduced the errors in the commercial DTM from 130 cm RMSE to 65 cm RMSE for the Enhanced DTM.

Phase 4 results are documented in two technical reports: "Verification and Validation Assessment of IFSAR Automated Feature Extraction" (June 2006) and "High Resolution Terrain Modeling Using Interferometric Synthetic Aperture Radar (IFSAR) Data" (June 2006).

In summary, this project used a disciplined approach to derive data collection requirements that are independent of system and vendor while yielding data to construct high-fidelity DTM, DSM and geospatial databases for FCS DTE and system-of-system M&S. Algorithms, methods, techniques, and software were developed into a critical toolset for processing LIDAR, IFSAR and photogrammetric data. Results of each project phase showed that this toolset provides a critical capability to the Army FCS program and other major defense acquisition programs.

The terrain enhancement and improvement methods described in this report rely on the use of control points obtained through ground surveys, precise LIDAR measurements, or other cadastral data, to remove uniformed and stochastic errors in IFSAR data. In denied areas, these control points are likely to be collected using LIDAR, differential GPS, or other remote sensing measurements taken by unmanned air and unmanned ground vehicles; unattended and persistent sensing; or dismounted special operations. While terrain enhancement is a function of the number of control points available, only a very small number of control points are needed to remove the large, uniformed biases in IFSAR-derived terrain models.

Table of Contents

List of Figures and Tables

Figures

Tables

Preface

The work was performed for Headquarters, U.S. Army Corps of Engineers (HQUSACE) by the Office of Technical Directors, Topographic Engineering Center (TEC). Technical review was done by Juan Perez, TEC's Systems Division Chief.

The TEC Principal Investigator was Associate Technical Director Dr. J. David Lashlee. Michael Coley is the Technical Director for Systems. The Deputy Director of TEC is Joseph F. Fontanella. The TEC Director is Robert W. Burkhardt.

TEC is an element of the U.S. Army Engineer Research and Development Center (ERDC), U.S. Army Corps of Engineers. The Commander and Executive Director of ERDC is COL Richard B. Jenkins, and the Director of ERDC is Dr. James R. Houston.

1 Introduction

Background

Test planning, rehearsal, and distributed test events for Future Combat Systems (FCS) require rapid generation of high-fidelity synthetic environments. These common environments consists of high resolution synthetic scenes of test and training ranges, which include the use of digital terrain models (both high and low resolution), digital surface models, 2-D and 3-D surface objects (man-made structures such as buildings, bridges, etc.) and other geospatial data necessary to accurately replicate conditions of each site.

By far the largest component of developing synthetic scenes in 3-dimensions is the commercial availability of **I**nter**f**erometric **S**ynthetic **A**perture **R**adar (IFSAR) and **LI**ght **D**etection **A**nd **R**anging (LIDAR) data collected from airborne platforms. Even though these industries are seeing rapid growth in the availability of data, in the numbers and types of sensors, and also in the vast amounts of data that can be collected on each flight mission with raw datasets exceeding hundreds of gigabytes in size. However the same type of growth cannot be seen with commercially available software in processing these types of data to provide high resolution, high accuracy topographic products. This is not to say this type of software does not exist, but there are evidently limitations in some of these products and in particular, in their ability to process and produce data quickly and accurately for FCS. This analysis of the current available IFSAR topographic data products was undertaken to identify these shortcomings in the industry, and to make recommendations to overcome them.

Objectives

The three main objectives of this project were to:

1. Assess the fidelity and quality of the commercial DSM and DTM from Intermap for the Yuma Wash area—Data Analysis.
2. Develop algorithms based on AFE for LIDAR data that can be applied to processing IFSAR data—specifically to perform foliage/vegetation removal and building/structure filtering while maintaining accurate terrain profile—Virtual Surfaces Method of processing IFSAR Data.
3. Assess improvements in constructing DSM/DTM using CCS-developed methods—Results from CCS IFSAR processing.

Approach

This study was conducted in a number of separate contracted efforts, which are summarized in the body of this report:

1. High Resolution Terrain Modeling Using Interferometric Synthetic Aperture Radar Data (Chapter 2)
2. Processing Color Imagery for White Sands Missile Range (Chapter 3)
3. Verification and Validation Assessment of Interferometric Synthetic Aperture Radar AFE (Chapter 4)
4. Light Detection and Ranging and Imagery Data Collection Requirements To Support the Generation of Synthetic Scenes and Digital Terrain Models (Chapter 5)
5. Generation of Synthetic Scenes and Digital Terrain Models Using LIDAR Data from White Sands Missile Range (Chapter 6).

Mode of Technology Transfer

This report will be made accessible through the World Wide Web (WWW) at URL: http://itl.erdc.usace.army.mil/library/publications.html

2 High Resolution Terrain Modeling Using Interferometric Synthetic Aperture Radar Data

Data Analysis

The analysis of the data consisted of four procedures:

1. Comparison of the Control Points against themselves (internal accuracy)
2. Comparison of the Control Points and 4m DEM as previously studied by U.S. Army Training and Doctrine Command (TRADOC)
3. Comparison of Control Points of Intermap IFSAR DSM and DTM
4. Comparison of DSM and DTM.

The following sections describe each procedure in detail.

Comparison of Control Points (Internal Accuracy)

In December 2004, Yuma Proving Grounds (YPG) funded the U.S. Army Training and Doctrine Command (TRADOC) Analysis Center-White Sands Missile Range (TRAC-WSMR) to conduct a field survey of an area of Yuma, Wash area to evaluate and validate a high resolution (4-m horizontal and 1-m vertical) DEM. The study area for this project was approximately 4 by 5 km, which was representative of the typical terrain for the whole of the original study area. GPS surveys were performed, which consisted of collecting points along 5 lines for a total of 2481 control points (Figure 1). To ensure data integrity for the current study, a series of tests were conducted to ensure the control points used for both studies exhibited similar results in that the internal accuracy of the GPS was correct. The results indicated that the control point data was of good quality, and well within acceptable survey limits for the purpose of this study for all lines (728, 729, 3665 and 3667) except for line (3666). The internal comparisons illustrated a bias of less than 10 cm for the acceptable lines and a bias of 92 -98 cm for the weaker line. This result was also similar to the conclusion as stated in the report TRAC_West_Fork_VV.pdf: "Four of the five survey lines required lowering the DEM about 0.3 m. However, line 3666 required the DEM be raised 0.8 m. No explanation could be found for this difference."

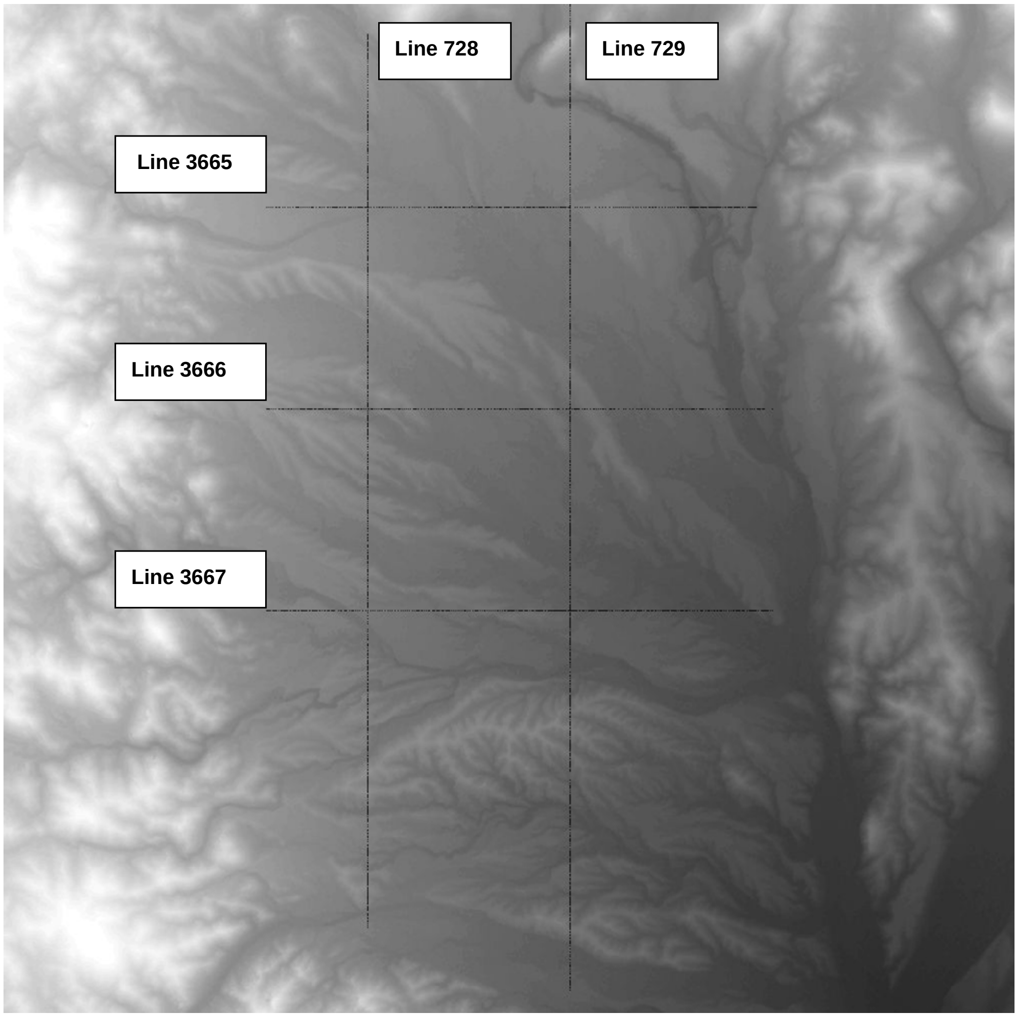


Figure 1. The dark lines represent the 2481 control points overlaid on the IFSAR DSM.

To validate the internal accuracy of the GPS control points, a comparison was made at the intersections of all lines. A z-value was computed for each line at the intersection using a linear interpolation. These two interpolated z-values were then compared against each other (Figure 2). Table 1 lists the equations of the differences between the intersections of lines.

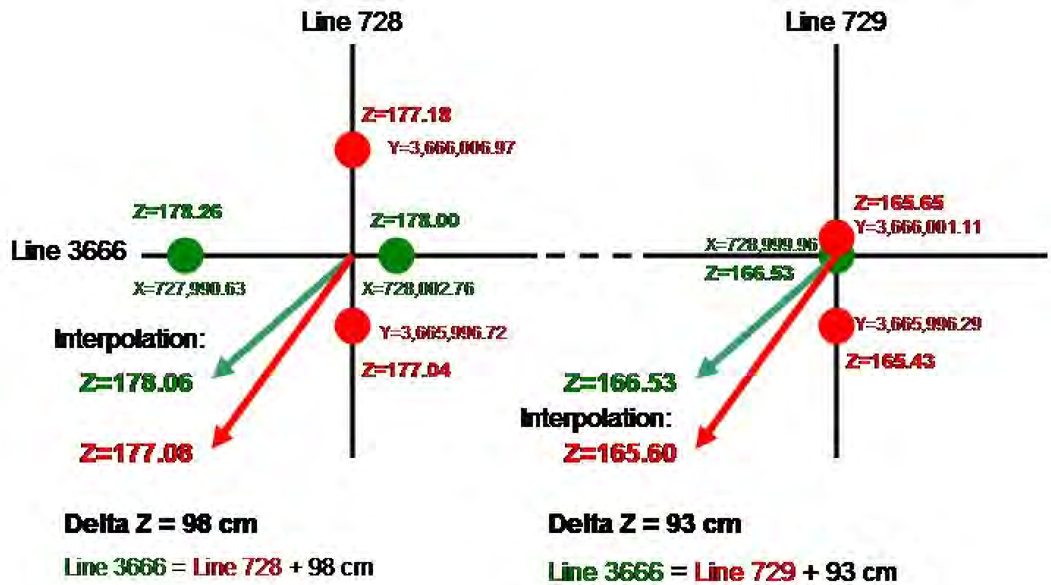


Figure 2. Illustrates the methodology used to evaluate the internal accuracy of the GPS control points by comparing the interpolated value at the intersection of the control point lines. By using an linear interpolation it can be seen that Line 3666 is higher than the line 728 and 729 by 98 and 93 cm respectively.

Table 1. Equations (1) illustrating the differences of the intersection of the control lines.

Computed Differences at Intersection of Control Point Lines
Line 3666 = Line 728 + 98 cm
Line 3666 = Line 729 + 93 cm
Line 3665 = Line 728 + 3 cm
Line 3665 = Line 729 – 10 cm
Line 3667 = Line 728 – 5 cm
Line 3667 = Line 729 + 4 cm

Comparison of Control Points and 4m DEM

To validate the results of Section 0 above, a comparison was made between the intersections of control point lines (Table 1) to the differences between the DEM and control points as derived by the “Field Validation of Yuma Wash Digital Elevation Model (DEM)” (WSMR 2005) (Table 2). Overall the comparisons exhibited good similarities where the differences between the first and second estimations (Table 3) are between 2 and 16 cm with an average of 9.7 cm. This concludes that control point lines agree with each other internally and externally using the 4m DEM. However this data also reiterates that a large shift still exists with the control points for line 3666, which constitutes 12 percent of the checkpoints. Therefore, 88 percent of the checkpoints are considered accurate.

Table 2. Illustrating the differences between the photogrammetrically derived 4m DEM and the control points.

Computed Differences between 4m DEM and Control Points
Line 728 = DEM (4m) – 31 cm
Line 729 = DEM (4m) – 24 cm
Line 3665 = DEM (4m) – 23 cm
Line 3666 = DEM (4m) + 80 cm
Line 3667 = DEM (4m) – 38 cm

Solving these equations yields a second estimation of the shift of heights between control lines (Table 3).

Table 3. Equations (2) illustrates the estimated error between the control point heights with the estimated error between the control points and DEM.

Second Estimation of Shift
Line 3666 = Line 728 + 111 cm
Line 3666 = Line 729 + 104 cm
Line 3665 = Line 728 + 8 cm
Line 3665 = Line 729 + 1 cm (2)
Line 3667 = Line 728 – 7 cm
Line 3667 = Line 729 – 12 cm

Comparison of Control Points to IFSAR DSM and DTM

By comparing the control points with the IFSAR DSM and DTM, a Root Mean Square Error$_{(Z)}$ is computed. All the values are high except for Line 3666, which has an RMSE of 73 cm. Yet, this line was identified as being suspect with a bias of approximately 1 m. The statistics in Table 4 therefore indicate a systematic bias in the DSM and DTM.

Table 4. RMSE and maximal errors between control points and the DSM and DTM.

RMSE and Maximal Errors Based on Control Points					
Control Point Line(s)	# of Pts	Surface Type	RMSE	Min Error	Max Error
All Lines	2481	DSM	117	-297	383
		DTM	126	-370	403
Line 728	684	DSM	118	-263	355
		DTM	125	-251	401
Line 729	736	DSM	124	-114	317
		DTM	137	-359	323
Line 3665	417	DSM	131	-103	383
		DTM	132	-125	430

RMSE and Maximal Errors Based on Control Points					
Control Point Line(s)	# of Pts	Surface Type	RMSE	Min Error	Max Error
Line 3666	295	DSM	73	-297	154
		DTM	90	-370	218
Line 3667	349	DSM	117	-154	357
		DTM	118	-271	386
4 Lines w/o 3666	2186	DSM	123	-263	383
		DTM	130	-359	403

To better understand the results, the distribution of errors (vertical distance between control points and surface – DSM or DTM) for the combined lines as well as for each separate line were plotted (Figures 3 to 8). By reviewing the peak of distributions of errors, the corresponding approximate elevation bias can be identified. For each line and their associated plots, a bias of approximately 1 m is evident except for line 3666 (Figure 7). This line illustrates a bias closer to zero with a lower RMSE than other lines. Again this line has been identified as an outlier, which proves that a bias does exist not only in this line, but also within the DSM.

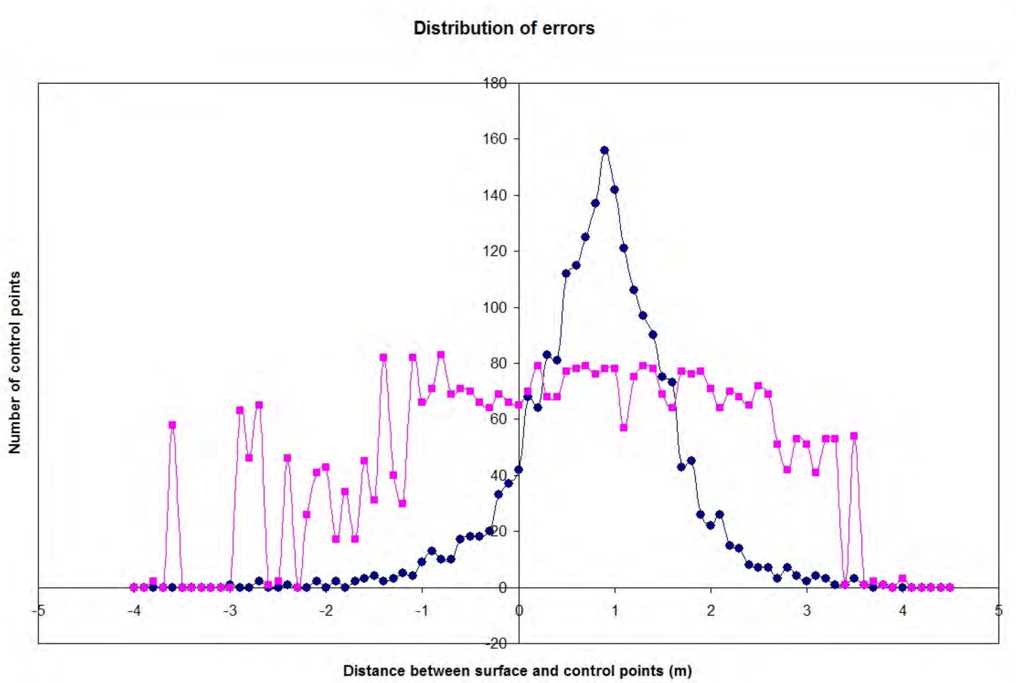


Figure 3. Distribution of errors (on 10 cm interval of z-distance) for all five lines of control points. Blue circles are errors of DSM (original data), red squares are DTM (bare Earth, processed by Intermap).

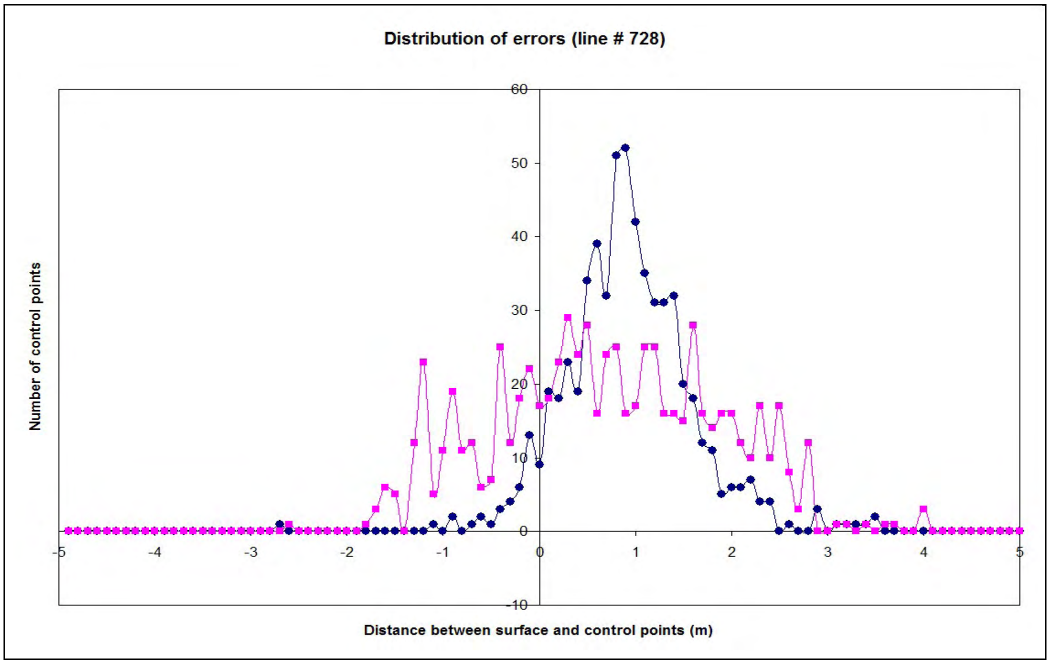


Figure 4. Distribution of errors for line No. 728. Blue circles are DSM, red squares are DTM.

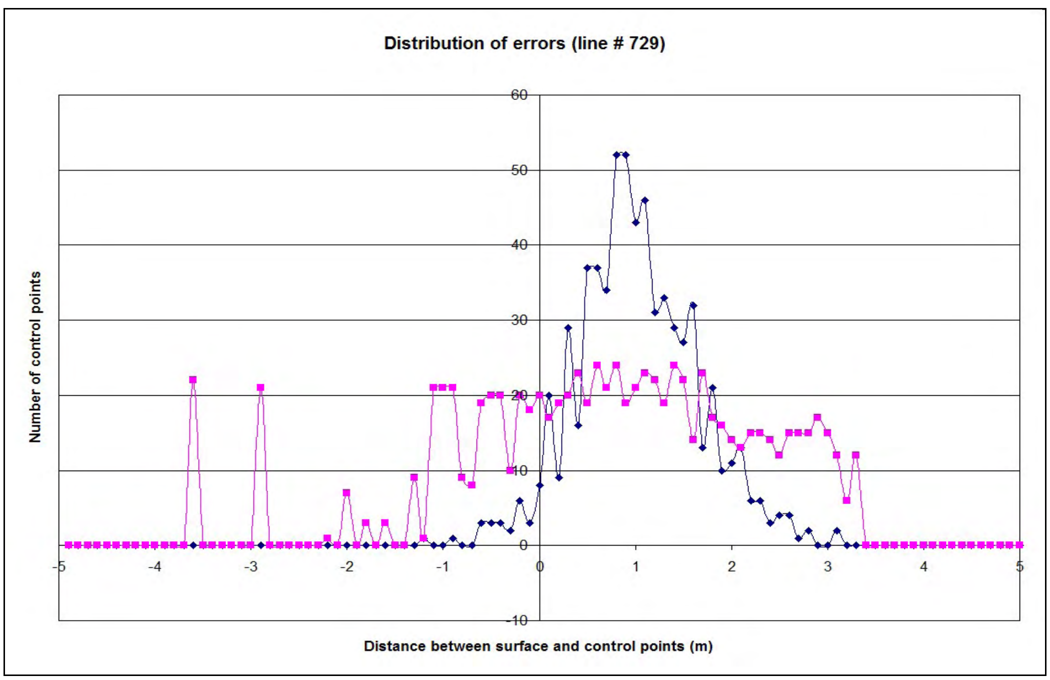


Figure 5. Distribution of errors for line No. 729. Blue circles are DSM, red squares are DTM.

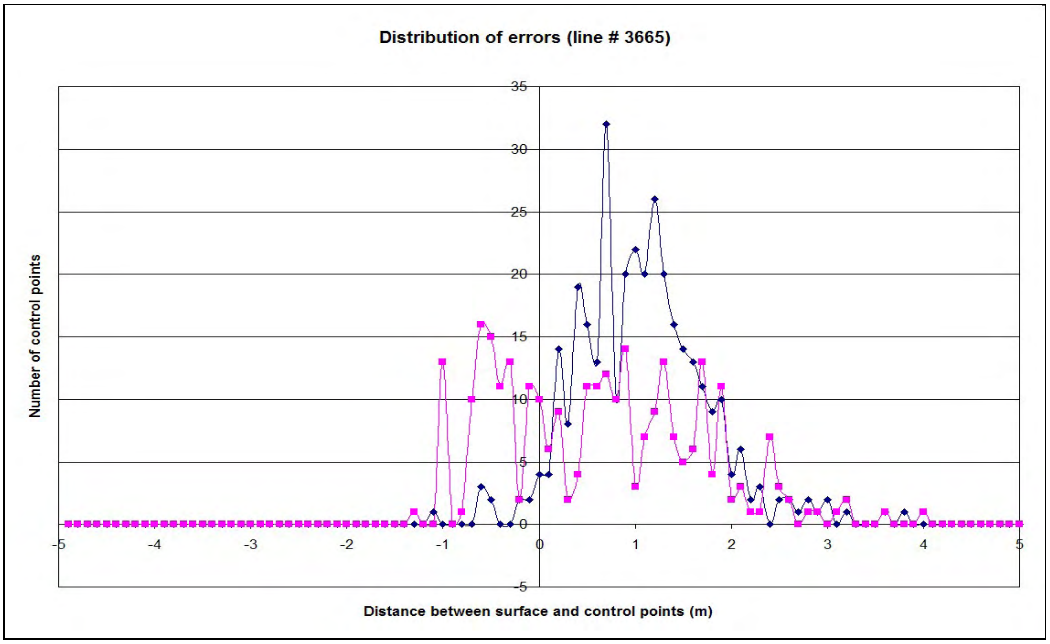


Figure 6. Distribution of errors for line No. 3665. Blue circles are DSM, red squares are DTM.

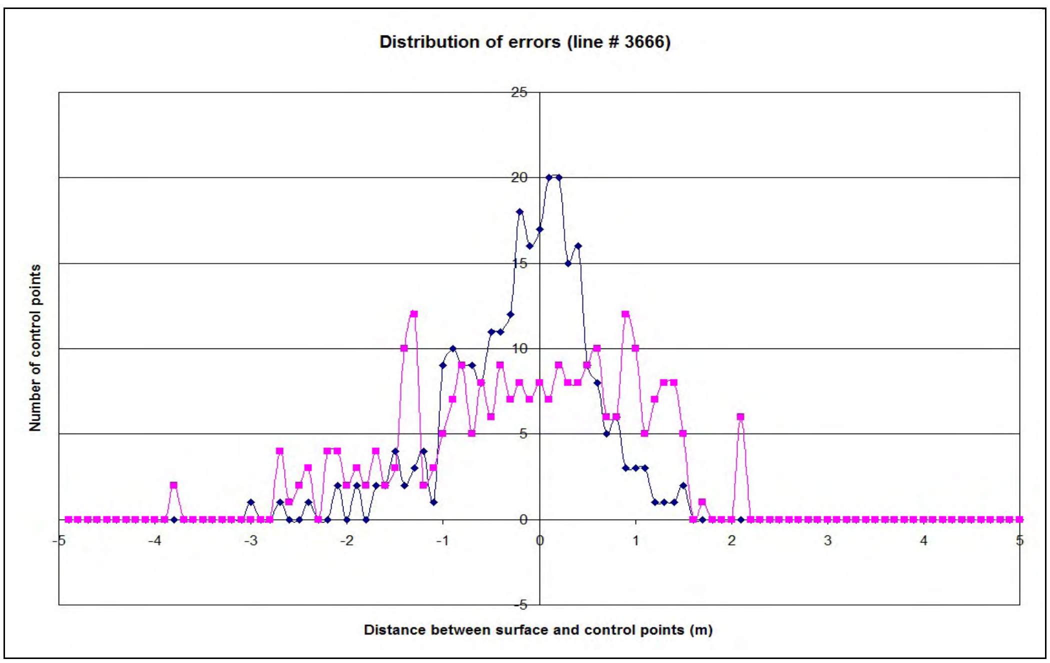


Figure 7. Distribution of errors for line No. 3666. Blue circles are DSM, red squares are DTM. Distribution of errors for DSM is different from other lines.

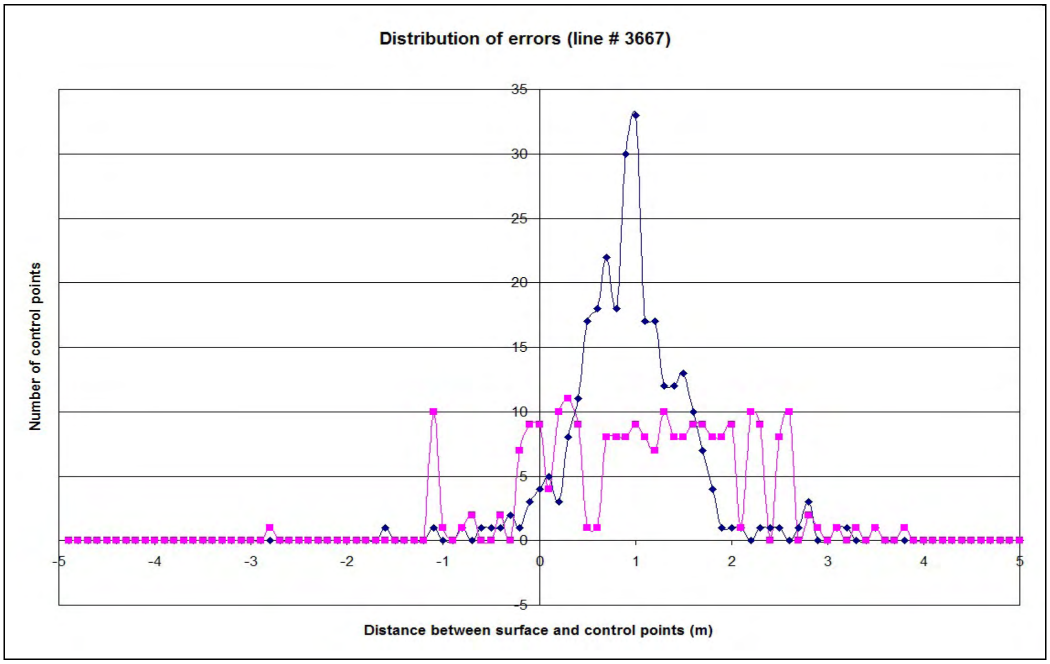


Figure 8. Distribution of errors for line No. 3667. Blue circles are DSM, red squares are DTM.

Comparison of DSM and DTM

A comparison of the actual model elevations to each other on a pixel-by pixel level, as opposed to just the control points, gives an additional perspective into the data quality. By using this methodology, one can see that control points are lower than either the DSM or DTM (Figures 9–10). Figure 11 shows that the control points from line 3666 appear to “fit” better, but they also show errors of up to 3m (Figure 12). Again, this line as been identified has being noncompliant for accuracy.

Additionally the DTM does not compare favorably to the DSM where it should. Since most of the DSM along the control lines are void of vegetation, it is conceivable that the DSM should equal the DTM and this clearly is not the case. It appears that the DTM is using a heavily interpolated process that is smoothed to derive the bare-earth product, and may not be representative of the true bare-earth elevations.

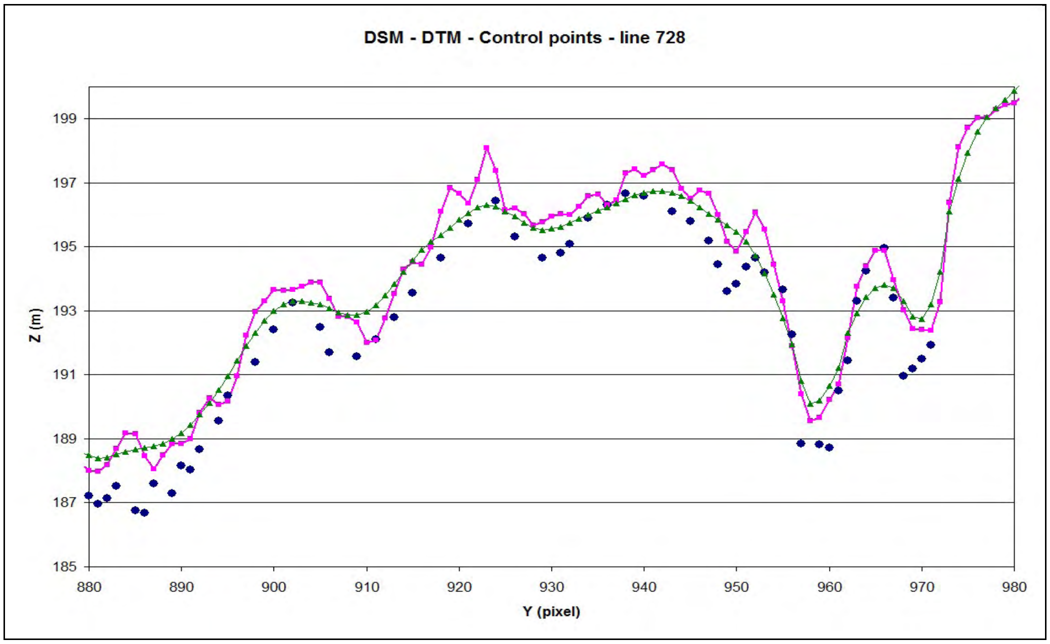


Figure 9. Intermap DSM, DTM and control points for 500 m (=100 pixels) of line No. 728. Blue circles are control points, red squares are DSM, green triangles are Intermap DTM.

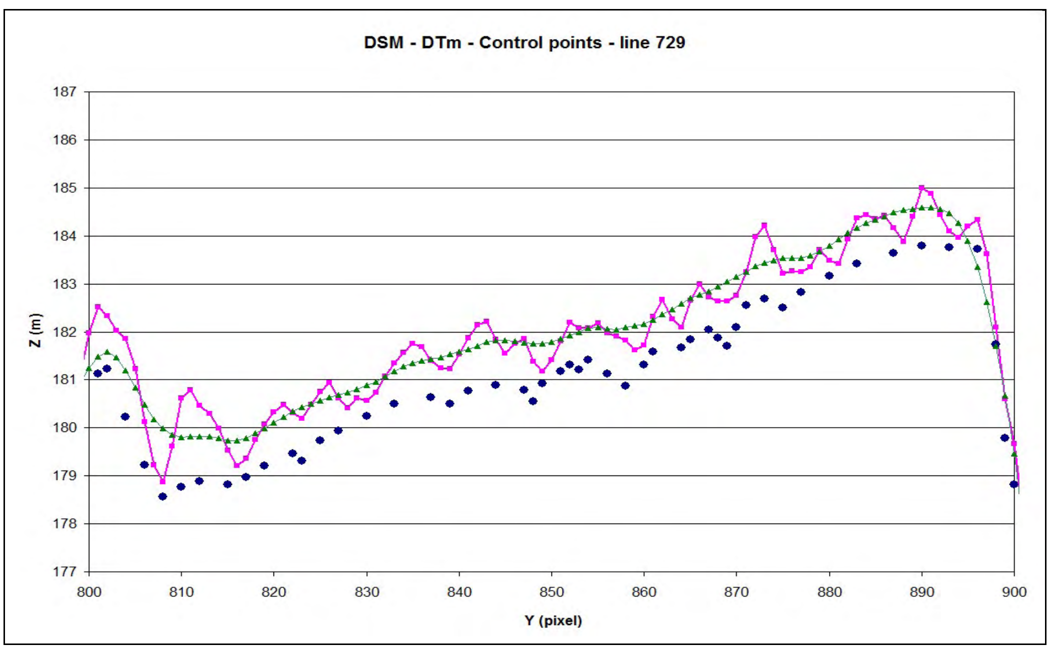


Figure 10. Intermap DSM, DTM and control points for 500 m of line No. 729. Blue circles are control points, red squares are DSM, green triangles are DTM.

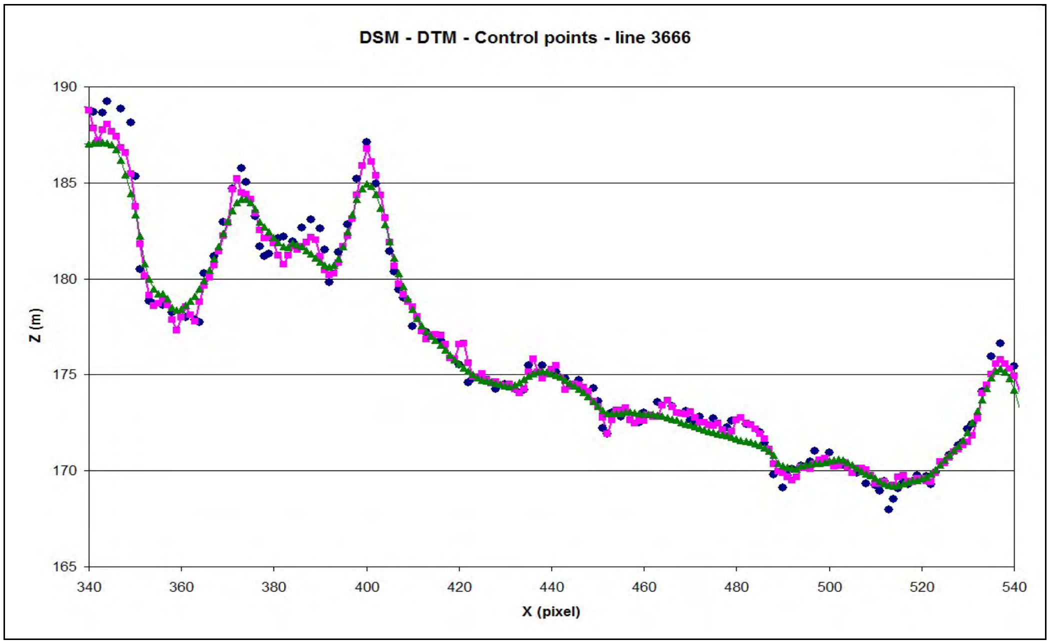


Figure 11. Intermap DSM, DTM and control points for 500 m of line No. 3666. Blue circles are control points, red squares are DSM, green triangles are DTM.

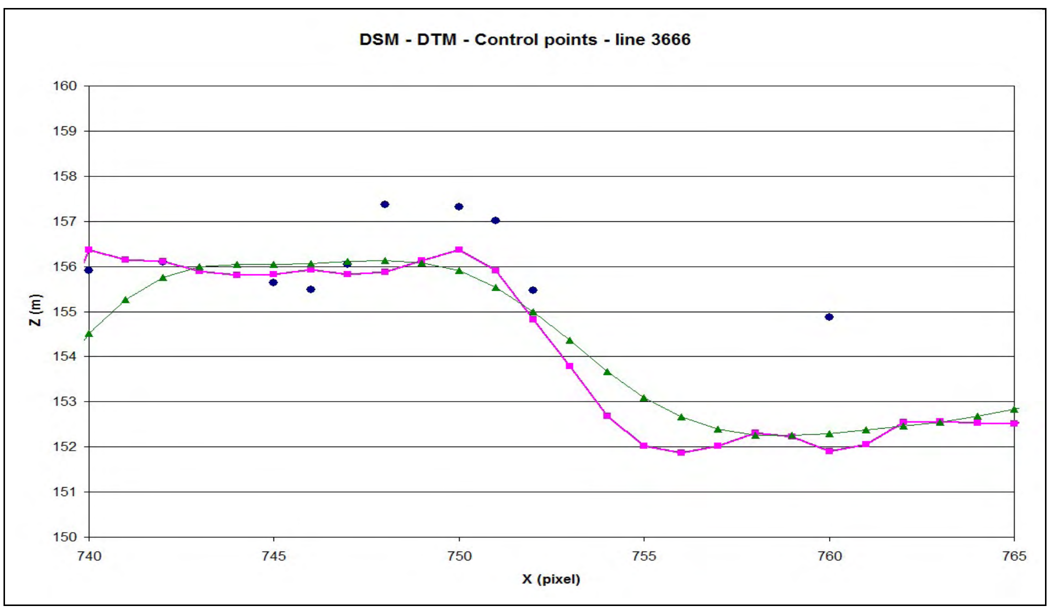


Figure 12. Intermap DSM, DTM and control points for 125 m (= 25 pixels) of line No. 3666. Blue circles are control points, red squares are DSM, green triangles are DTM. Right control point show large (~ 3m) shift from DSM.

Virtual Surfaces Method of Processing IFSAR Data (2)

CCS algorithms that include Virtual Surfaces Method for Automatic Feature Extraction and bare-earth (DTM) generation were developed, tested, and used for processing thousands of square km of LIDAR data throughout North America. Using the same principles and approach, these algorithms were modified to incorporate the characteristics of IFSAR data to produce a superior DTM that reflects the true terrain elevations with a better surface estimation. Results of this project show that the CCS method for DTM generation is significantly increased for not only the absolute accuracies, but also the relative accuracies of the DTM. This methodology produces a bare-earth DTM that is more representative of the true terrain values and is not over-smoothed typical of the Intermap DTM.

Basic Principles of Virtual Surface Method (VSM) for Processing 3D Data.

Using the DSM as the source, the basic principle of this VSM for processing IFSAR/LIDAR data is to generate a virtual synthetic surface, which will emulate the bare-earth terrain, but without vegetation or artificial structures such as buildings. This method uses the complex principles of spherical surface estimation.

Figure 13a shows the DSM as a dashed line in which a spherical segment surface must be generated using the minimum elevations of the terrain. Figure 13b shows the virtual surface from a set of top surfaces of spherical segments. As can be seen from Figure 13b, the virtual surface follows the DSM well excluding the area containing the building. For the area of the building, a sphere is created in the location of the extracted feature (Figure 14a) as well as a mirror sphere (Figure 14b). This surface differs from Figure B in areas of extracted objects and deep hollows. From these products an average surface is created (Figure 15a), but in the areas of the hollows, the original minimum elevations must be retained. In these areas, the elevations from Figure 13b are re-used. The end result is Figure 15b, which maintains the integrity of the DSM, but does not contain extraneous artifacts (providing a superior surface estimation).

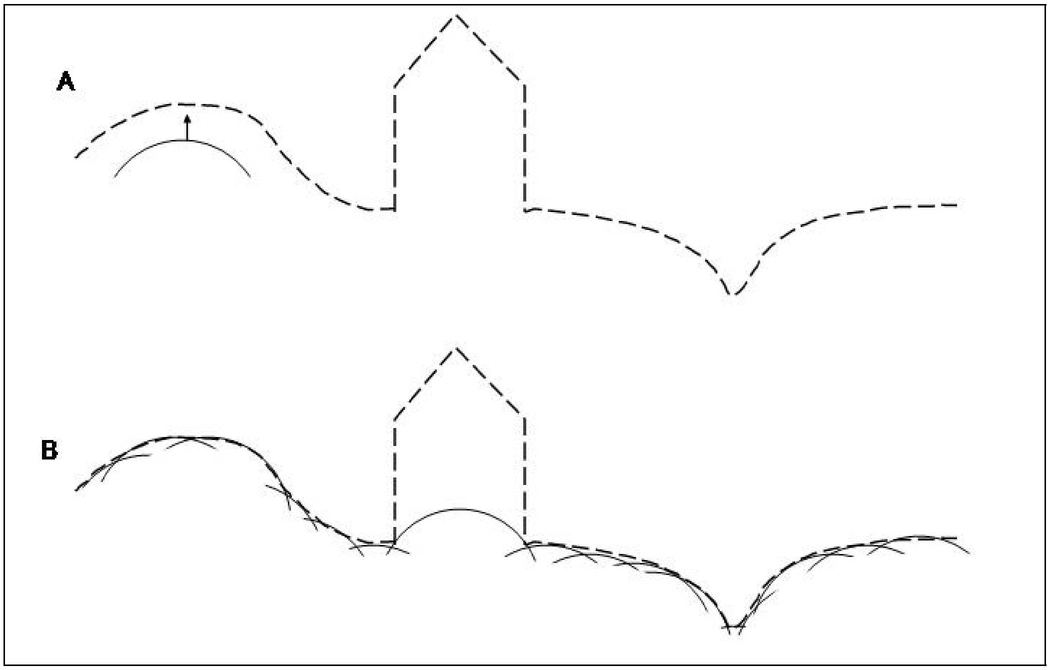


Figure 13. Generation of different surfaces within the Virtual Surfaces Method (VSM). A – original DSM, B – Surface from spherical segments.

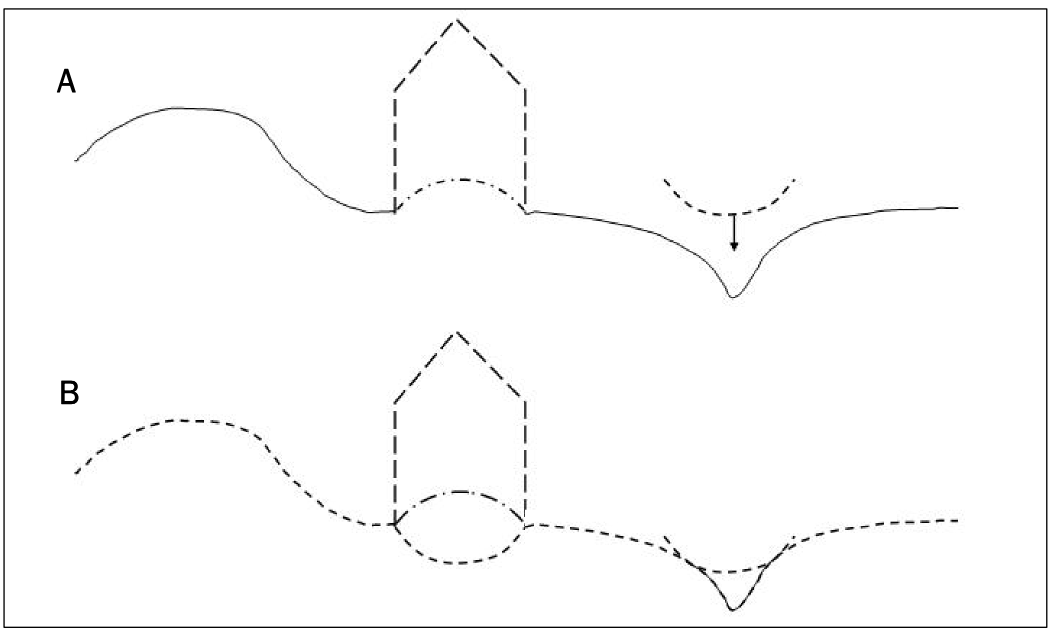


Figure 14. Generation of different surfaces within the Virtual Surfaces Method (VSM). A – Spherical estimation of extracted features and deep hollows, B – Surface from mirror spheres of extracted features.

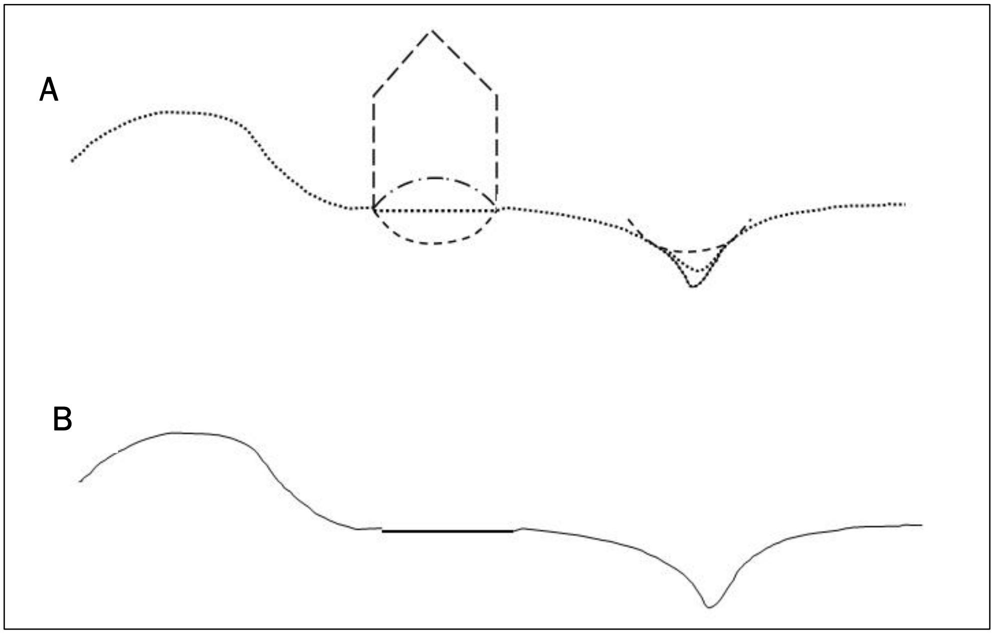


Figure 15. Generation of different surfaces within the Virtual Surfaces Method (VSM). A – Average surface estimation, B- Average surface estimation while retaining low terrain points (hollows).

Figure 16 shows the process flow of the product generation. The power of this methodology is its ability to use these automated approaches, thereby minimizing the manual verification processing. This method would support on-board processing in the event of real time needs.

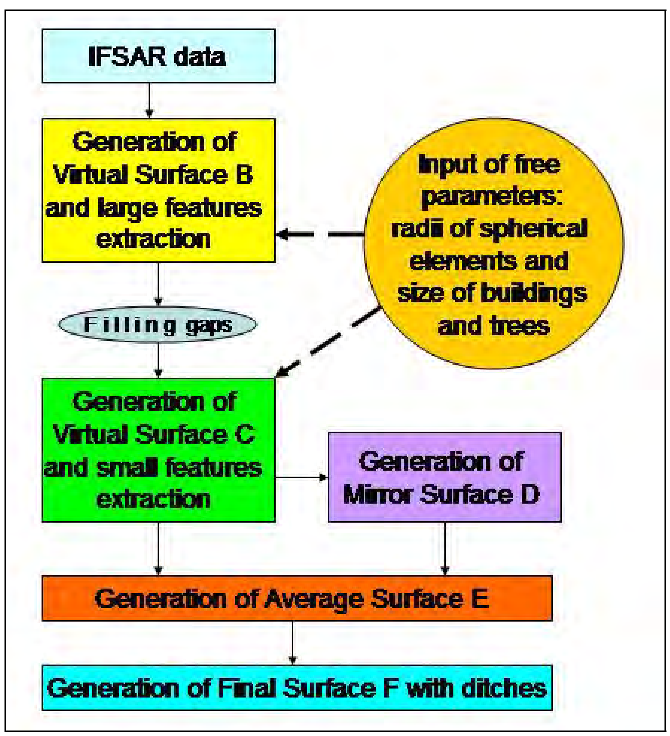


Figure 16. Algorithm process flow of IFSAR data for this project.

Results from CCS IFSAR Processing (3)

Benefits of AFE Methods by CCS vs. Methods by Intermap.

Since the control points from Line 3666 do not illustrate a systematic shift (even though it does exist, as both the control points and DTM include a bias), this can be used as a baseline to compare the DTM from Intermap to the DTM from CCS. Figures 17 to 21 show that the CCS processed DTM better fits the control points. The RMSE for this line improves from 90 cm for the Intermap DTM to 73 cm for the CCS DSM. Additionally the minimum and maximum errors are reduced by over 60 cm. With more accurate control point, these values could improve further. It is clear that the Intermap DTM, although esthetically pleasing, is not the best representation of the true digital terrain elevations. The results also show that this Automatic Feature Extraction (AFE) and method of DTM generation provides a balance of speed of processing while maintaining the integrity of the data.

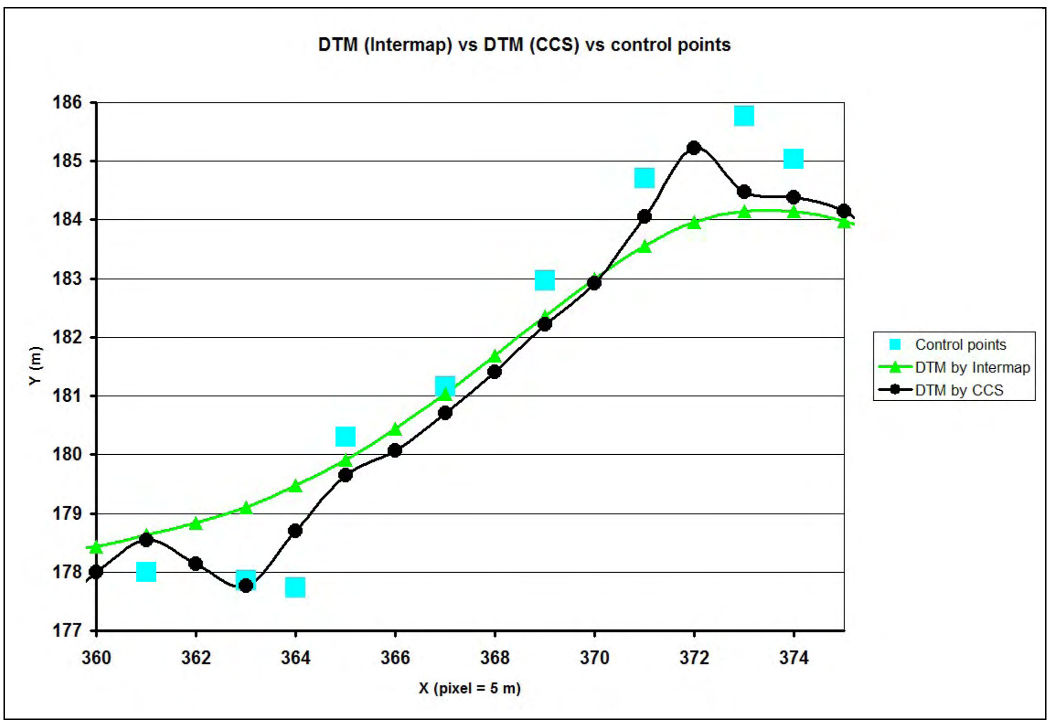


Figure 17. Comparison of Intermap DTM, CCS DTM, and control points for 75 m (= 15 pixels) of line No. 3666. DTM by Intermap contains over-smoothed hills and hollows. DTM by CCS follows control points better.

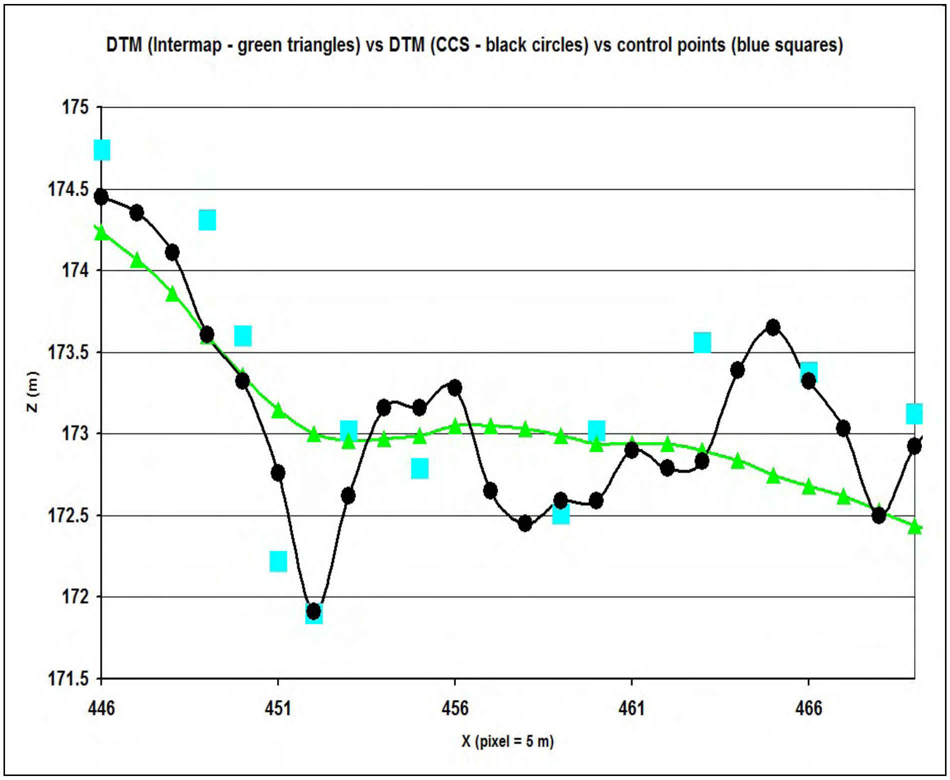


Figure 18. Comparison of Intermap DTM, CCS DTM, and control points for 65 m (= 13 pixels) of line No. 3666. DTM by Intermap DTM ignores small details of relief. DTM by CCS follows control points better.

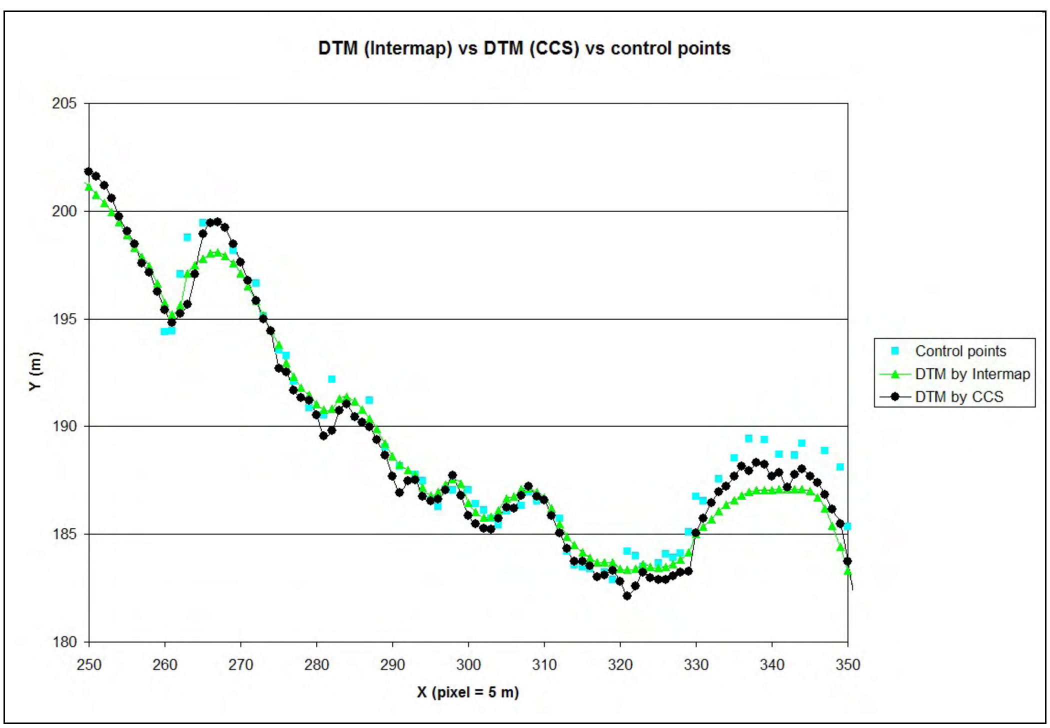


Figure 19. Comparison of Intermap DTM, CCS DTM, and control points for 500 m (= 100 pixels) of line No. 3666. DTM by CCS follows control points better.

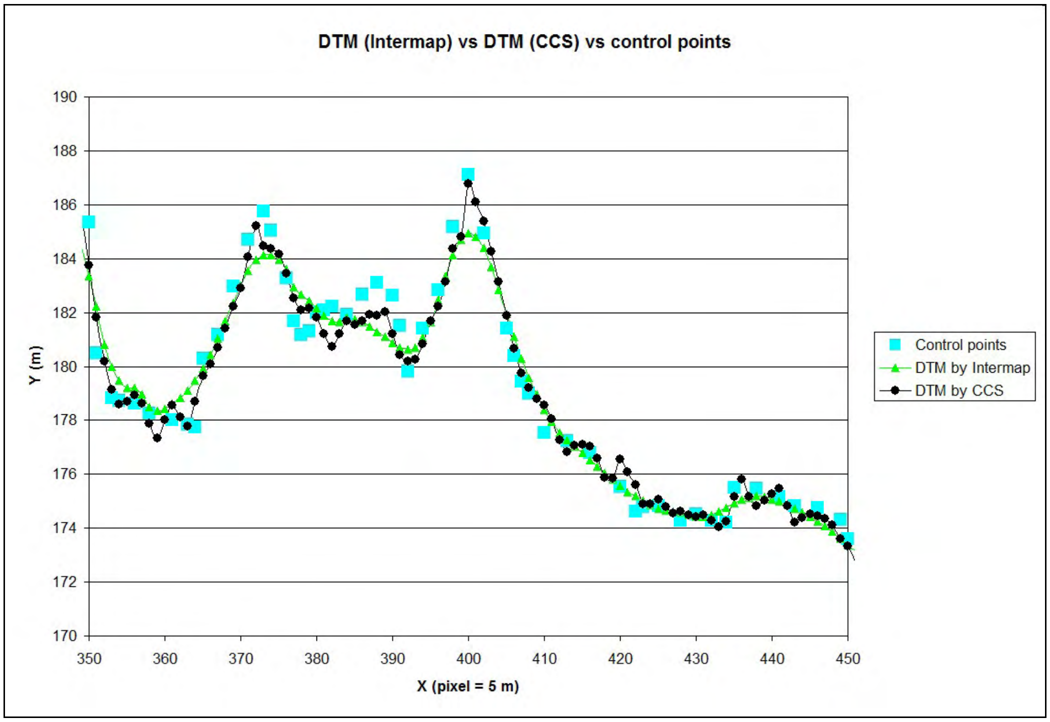


Figure 20. Comparison of Intermap DTM, CCS DTM, and control points for next 500 m (= 100 pixels) of line No. 3666.

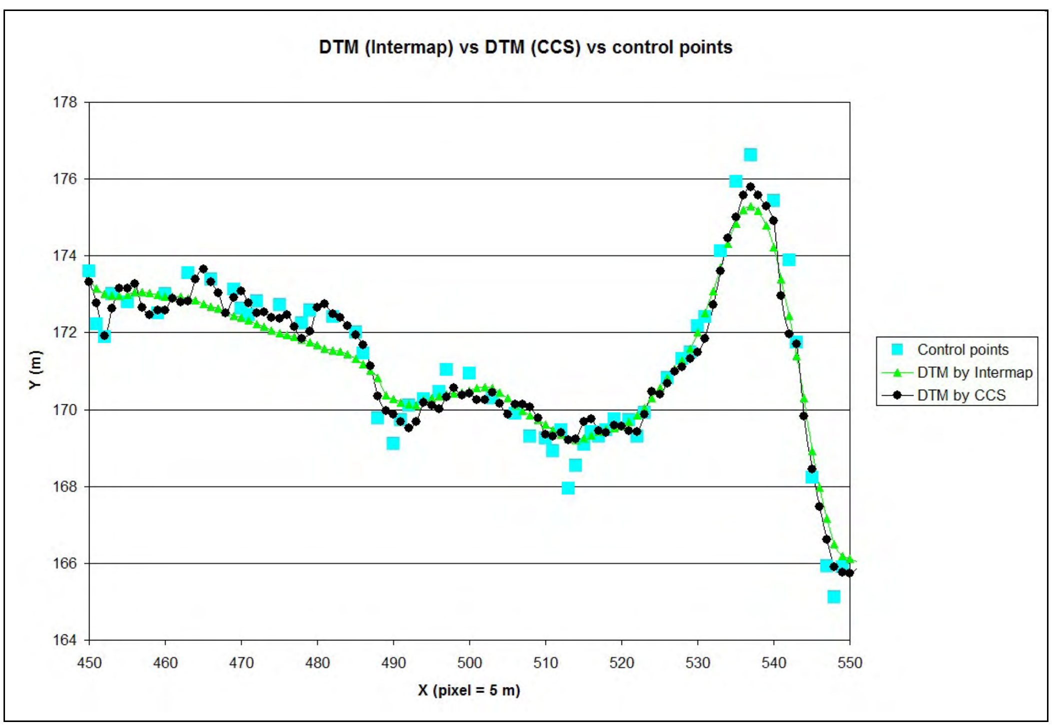


Figure 21. Comparison of Intermap DTM, CCS DTM, and control points for next 500 m (= 100 pixels) of line No. 3666. Within 1,5 km of line of control points Intermap DTM has similar behavior, but ignores small details of relief with a height or a depth of ~1–2 m.

Improving DTM after V&V analyses.

Using Verification and Validation (V&V) methods and analyses revealed that the original DSM and DTM by Intermap exhibit a systematic bias (Z-shift) from control points in lines No. 728,729, 3665, 3667 (~88 percent all control points). It was also found that accuracy of DTM improves if the Intermap data is lowered by 1 m. Figures 22 to 26 show results of DTM generation with this correction. Difference between DTM by Intermap and DTM by CCS is the result of different processing approaches for DTM generation and vertical correction after V&V process. The end results show that together this AFE algorithms and V&V method can make drastic improvements in the Intermap DTM. For all 2481 control points in five lines:

Intermap DTM RMSE = 126 cm with average error of 104 cm and maximal errors between -370 cm and 403 cm.

CCS DTM RMSE = 77 cm with average error 57 cm and maximal errors between 397 cm and 283 cm.

After ignoring possible wrong control points in line No. 3666 (295 points) and using 2186 control points, Intermap DTM RMSE = 130 cm (with average error 109 cm and maximal errors –359 cm and 403 cm), and CCS DTM RMSE = 65 cm (with average error 49 cm and maximal errors –363 cm and 283 cm). In all cases, the data improves not only statistically, but the overall data integrity of the true surface elevations is maintained.

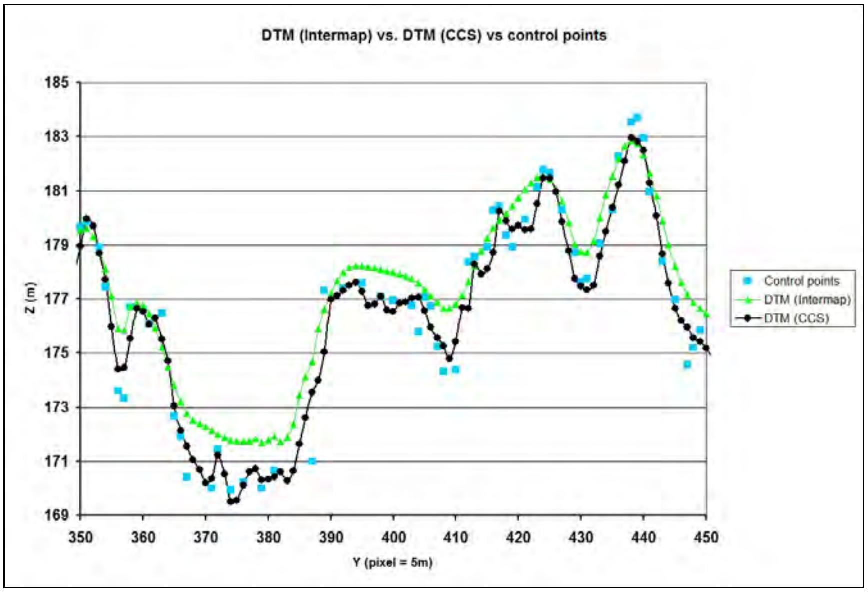


Figure 22. Comparison of Intermap DTM, CCS DTM and control points for 500 m (= 100 pixels) of line No. 728. DTM by CCS follows control points much better. DTM by Intermap is not only over-smoothed, but contains a 1 m bias (z-shift) above this DTM and control points. Typical difference in heights between two DTMs is ~ 1–2 m.

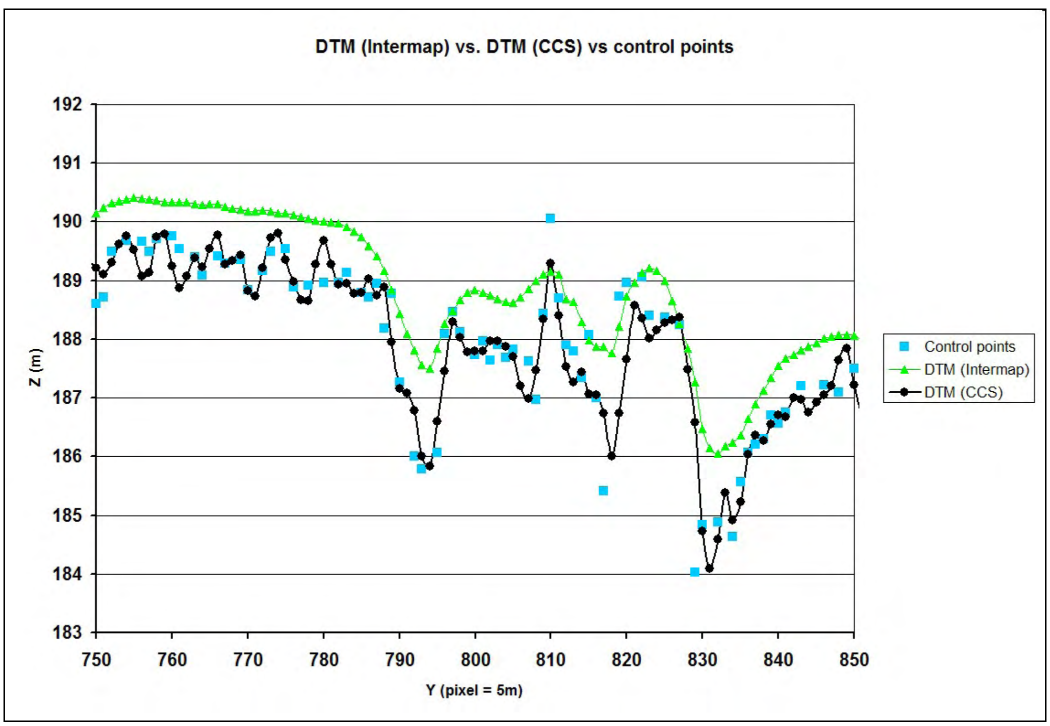


Figure 23. Comparison of Intermap DTM, CCS DTM, and control points for 500 m (= 100 pixels) of line No. 728. DTM by Intermap is over-smoothed and shifted above this DTM and control points.

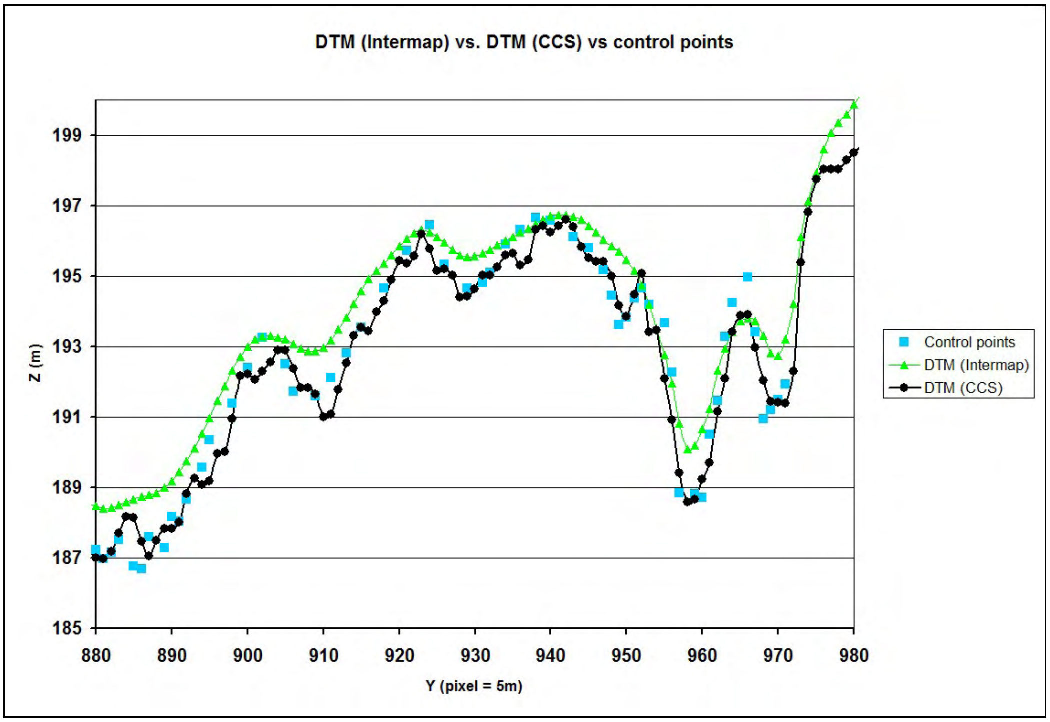


Figure 24. Comparison of Intermap DTM, CCS DTM, and control points for 500 m (= 100 pixels) of line No. 728. DTM by CCS follows control points much better.

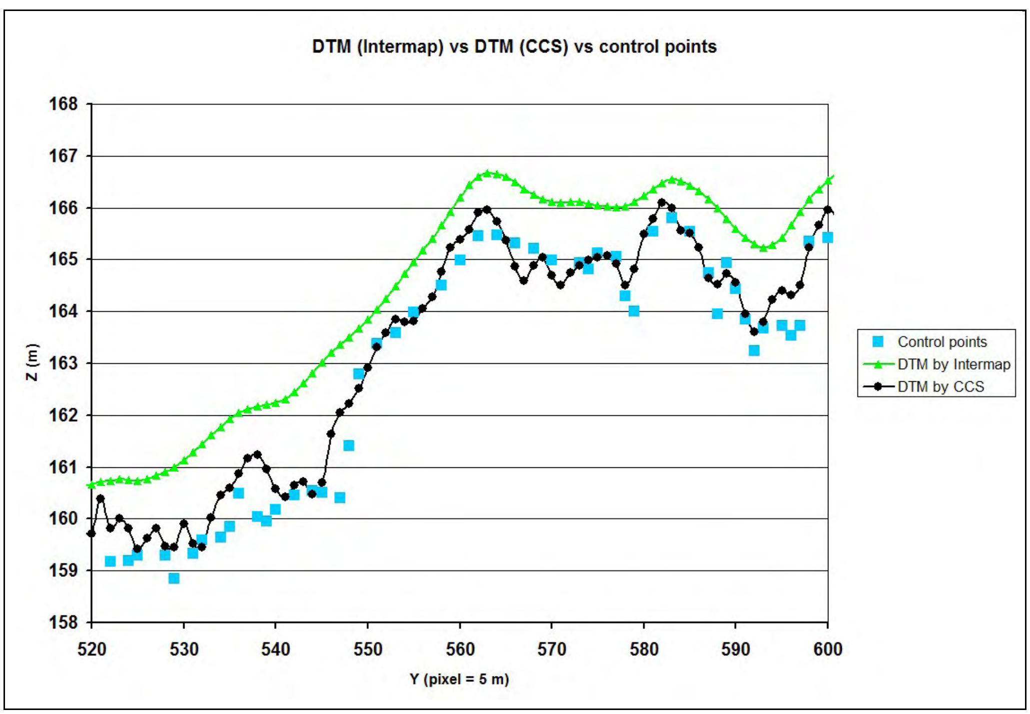


Figure 25. Comparison of Intermap DTM, CCS DTM, and control points for 500 m (= 100 pixels) of line No. 729. DTM by CCS and DTM by Intermap are drastically different. Differences in heights between two the DTMs reached 2–2.5 m.

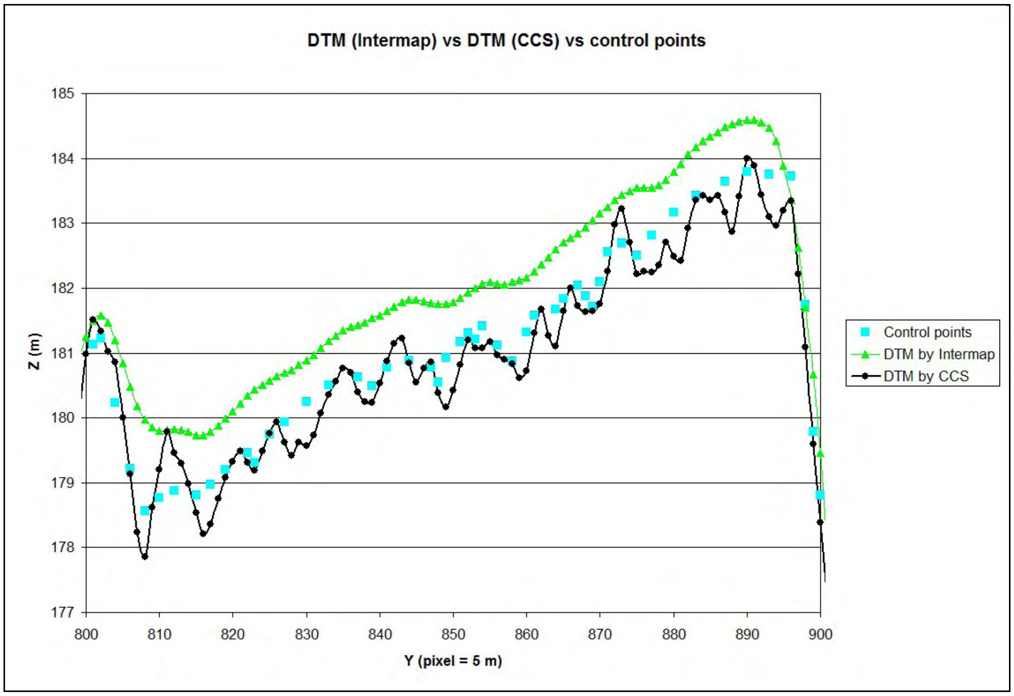


Figure 26. Comparison of Intermap DTM, CCS DTM, and control points for 500 m (= 100 pixels) of line No. 729. DTM by CCS follows control points better.

Conclusion

It is clear from the data that IFSAR can be a viable technology for producing topographic data products for synthetic modeling applications. It is capable of acceptable accuracies in areas of minimal vegetation, but has limitations in areas of heavy vegetation. For this data study, two types of errors existed: the absolute accuracy, and the originally processed DTM, which consisted of an over-smoothed, and highly interpolated surface, not representative of the terrain. By using good control points, the overall absolute accuracy can be improved. In this case, shifting the data 1 m made a significant improvement. The additional improvement can be obtained from processing the DTM from the DSM using the CCS methods as described in this report.

This work concludes that:

1. The Intermap DSM provides a relatively good representation of the surface; however, the DSM can be improved using the CCS techniques and V&V measurements for removing errors and biases.
2. Analysis of the Intermap DSM to the 2,481 control points of the Government Furnished Data (GFD) showed an elevation accuracy of 1.17 m RMSE.

3. Analysis of the Intermap DTM to the 2,481 GFD control points showed elevation accuracy of 1.26 m RMSE.
4. Further analysis showed that the Intermap DTM is further degraded with increase foliage density and terrain gradient or offsets.
5. Results show that the CCS AFE techniques and other terrain/surface construction methods can significantly improve the accuracy of the IFSAR derived DTM. The CCS team was able to successfully remove surface objects and foliage without adversely affecting spatial accuracy. Furthermore, the CCS methods are expected to provide substantial improvements to IFSAR derived DTMs in areas of heavy vegetation.
6. Across the set of control points, implementation of the CCS algorithms reduced the errors in the DTM from 35 to 50 percent. For 2186 control points, the Intermap DTM has an error 130 cm RMSE, and CCS DTM has an error of 65 cm RMSE.
7. Control points are necessary for generation of high quality DTM from IFSAR data, and a procedure for needs to be developed to optimize the V&V process.

3 Processing Color Imagery for White Sands Missile Range

Description of Data Source

Map and Geographical Type of Project Area.

The project area was a 60 x60 km parcel, in White Sands, Otero County, NM. The map in Figure 27 (in red) shows the area of interest, covered partially by imagery in ~50x25 km region (each tile has a size ~6x7 km (Figure 28). WSMR is a flat desert area with bushes and grass and small number of buildings (exclude one village).

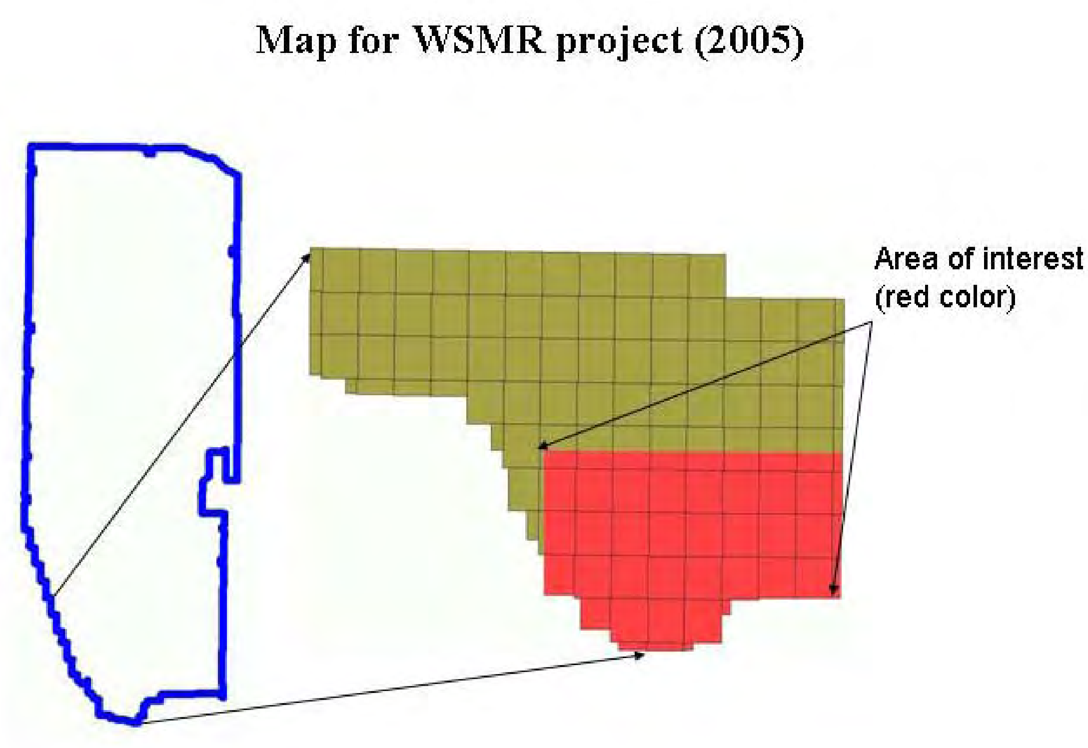


Figure 27. The area that covered by color imagery is outlined in blue. The project area is red.

Figure 28. Area 6x7 km chosen for pilot project.

General Description of Data Products for White Sands Missile Range Project

The type of data and size of the tiles for White Sands Missile Range Project was:

- color imagery (R,G,B) with 1-ft resolution
- each image covers ~ 6 x 7 km tile (18,000 x 21,000 pixels) and has ~1 GB size.

The large tile/image 32106D44 was split into sub-tiles 300x300 m (1,000x1,000 pixels) without overlap (Figure 29).

Figures 30 to 32 illustrate three subtiles 300x300 m with various landscape characteristics found in the project area (desert with road, desert with old river channel, urban area). Generally the project area is a sand (white, yellow, or grey) desert, covered by bushes and tall grass.

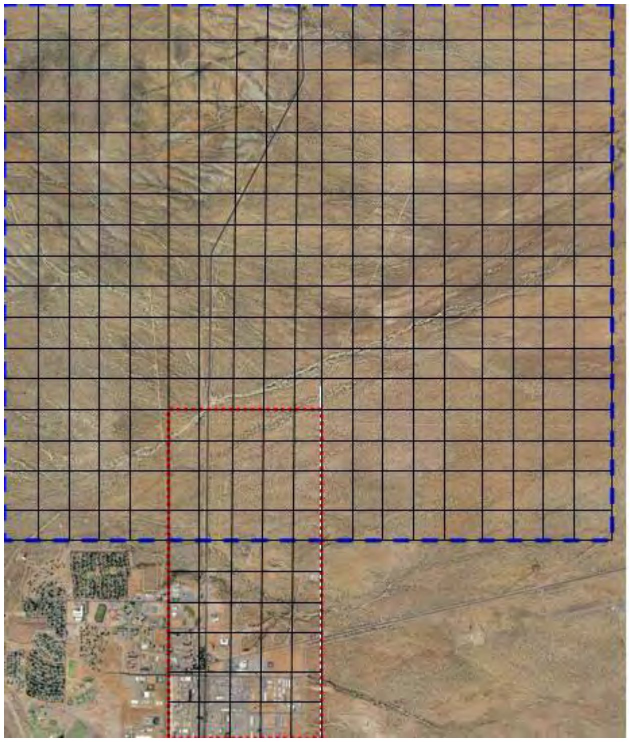


Figure 29. Pilot project area was split into tiles 300x300 m size (1,000x1,000 pixels). 323 tiles inside the blue line was processed for foliage extraction. 50 tiles inside the red line were used for roads polyline generation

Figure 30. Grey desert with dark foliage and asphalt road.

Figure 31. Yellow desert with green foliage, unpaved road, trail and dry river channel.

Figure 32. Urban area.

Spectral Characteristics of Collected Color Imagery

An important characteristic of the collected color imagery is differences of spectral parameters of light desert soil and dark foliage (bushes and tall grass). Figures 33 to 36 show typical cases that contribute to the analysis of spectral variability of desert and vegetation.

Figure 33. Yellow desert with minimum foliage.

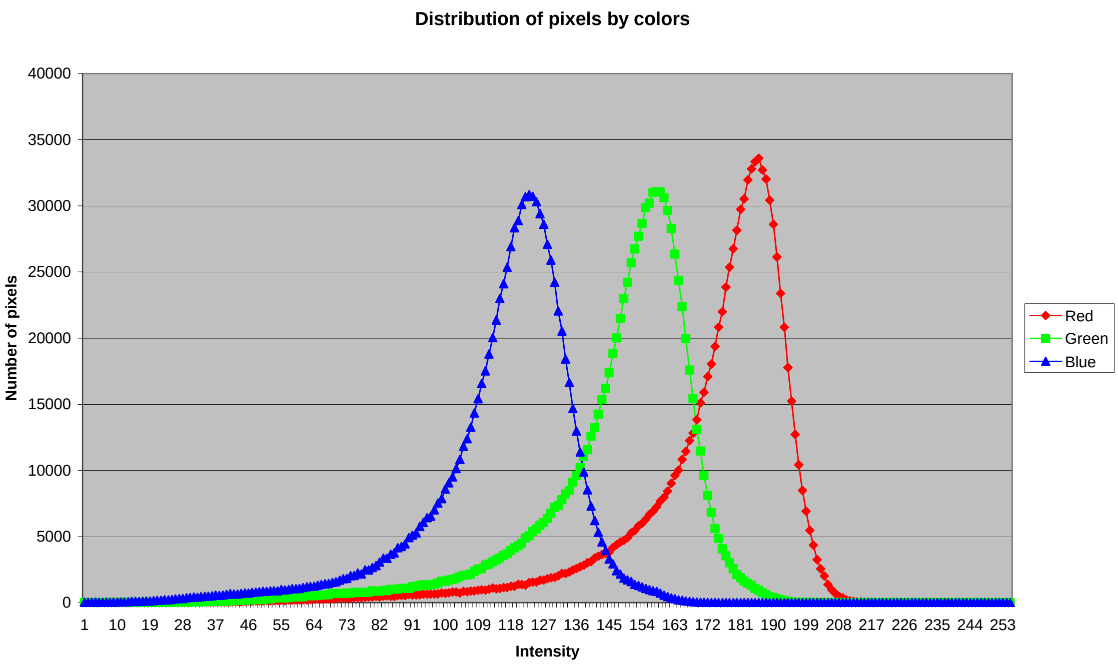


Figure 34. Distribution of pixels by colors for yellow desert with minimum foliage (see Fig. 31). Red is dominate, blue is minimal. General brightness of desert is ~150 (in scale between zero and 255)

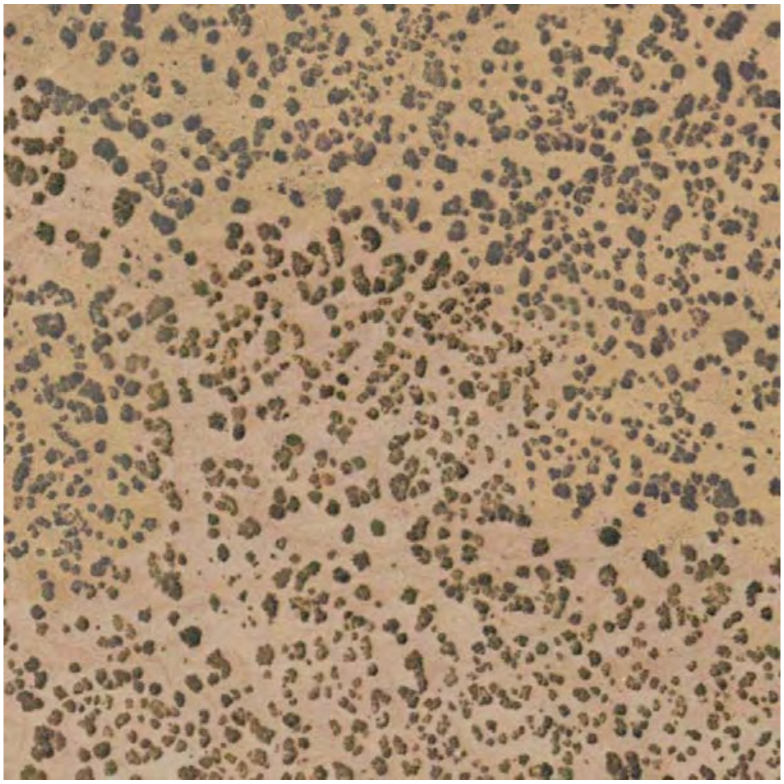


Figure 35. Yellow desert with numerous bushes with high brightness.

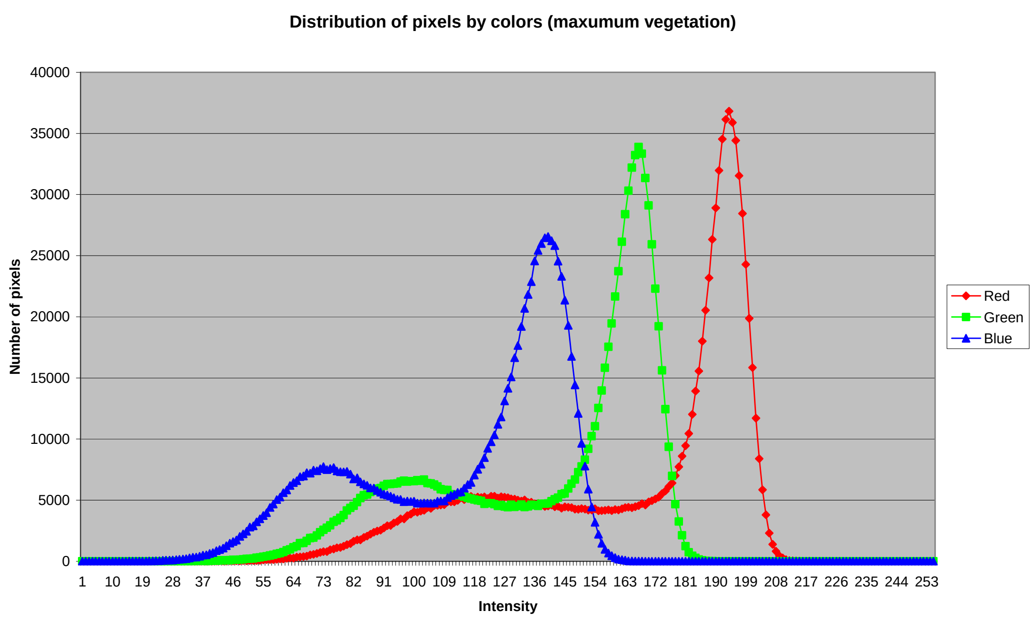


Figure 36. Distribution of pixels by colors for yellow desert with maximum foliage (see Fig. 35). Small maximum in curves on the left associated with foliage with average high brightness ~60–120.

Figures 37 and 38 show some of the variability of spectral characteristics resulting from different weather conditions and some errors during collection of imagery.

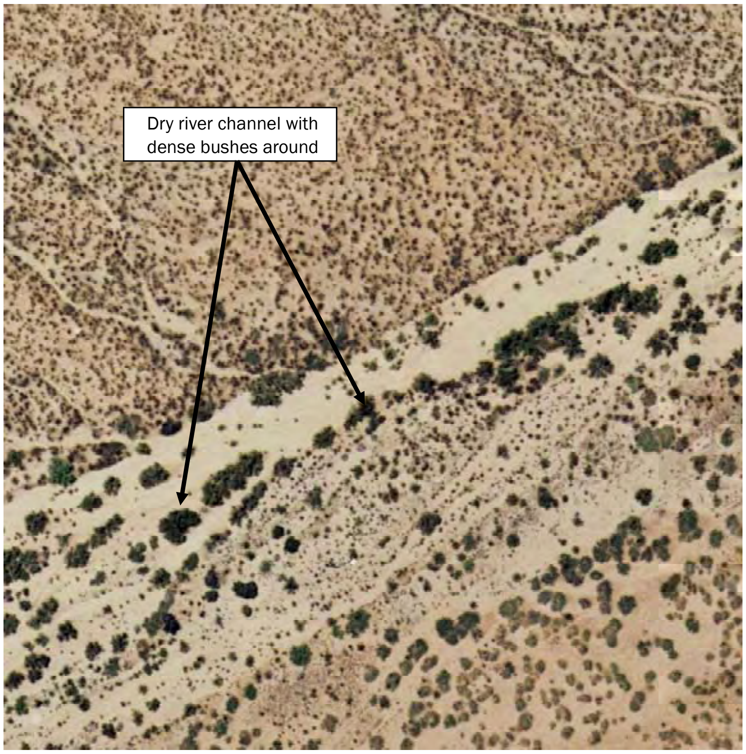


Figure 37. Dry river channel with white sand and different types of desert foliage around (large bright bushes in south, small dark foliage in north).

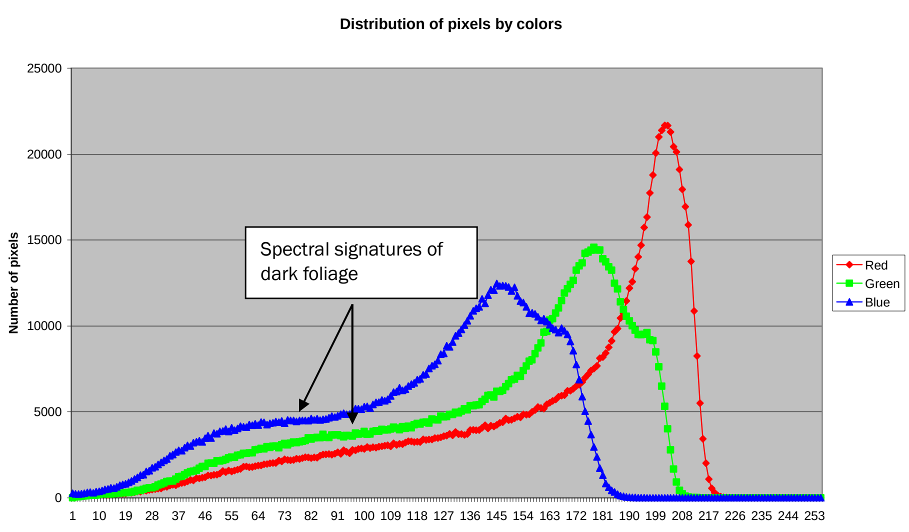


Figure 38. Distribution of pixels by colors for yellow desert with mixed types of foliage (cf. Fig. 37). Spectral signatures of foliage is different and cover area of brightness from 30 to 120.

Vegetation Modeling

For extraction of foliage, pixels are classified as:

1. Vegetation and as desert.

Classified Imagery Database

According to this analysis, the next types of pixels that can be separated in the imagery are:

2. Pixels with high reflectivity (~150) and domination of red color that represent dry desert soil
3. Pixels with medium or low reflectivity (<120) and domination of green color that represent part of desert foliage
4. Pixels with medium reflectivity (~100–120) and domination of red color or grey color (quasi-equality of different colors) that represent transparent part of desert foliage (each pixel consist both foliage and soil)
5. Pixels with low reflectivity (~80–100) and red or grey color that represent dark part of desert foliage
6. Pixels with very low reflectivity (<80) and red or grey color that represent darkest part of desert foliage
7. Dark pixels with very low reflectivity (<10) and grey color that represent mostly shadows.

Figure 39 shows the theory of extraction of vegetation as a set of pixels of two type of vegetation: dark foliage (black) and light green vegetation (grey). Clustering of vegetation is performed by separation of sets of pixels, which are connected by the sides only, not corners.

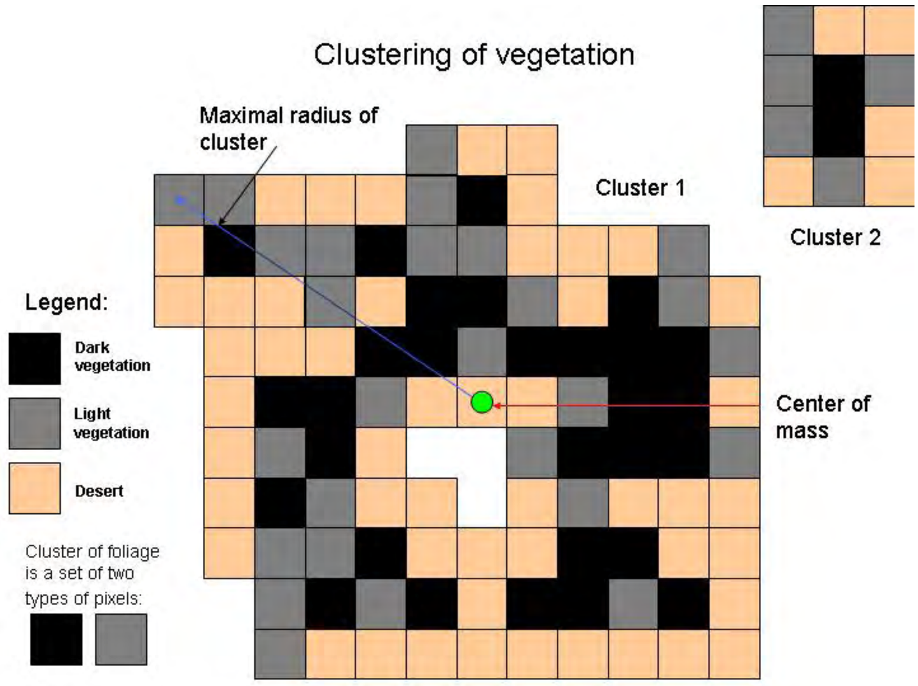


Figure 39. Clustering of vegetation.

Each foliage object in the foliage database is a cluster (a set of pixels, connected by sides, not by corners) from next types of pixels:

8. Very dark pixels identified as vegetation:

   Criteria: I_green<80. Flag (value of pixel) = 3.

9. Dark green pixels identified as vegetation:

   Criteria: I_green>I_red and I_green>I_blue and I_green<120.
   Flag (value of pixel) = 2.

Due to non-uniformity of parameters of foliage, for 21 percent tiles (69 from 323), an additional class of pixels was considered as vegetation:

10. Medium dark pixels:

    Criteria: 80<I_green<100. Flag (value of pixel) = 4

For tiles with relatively bright soils and vegetations (7 from 323) another additional class of pixels was considered as vegetation:

11. Light dark pixels:

    Criteria: 100<I_green<120. Flag (value of pixel) = 5

From this raw estimation, more than 90 to 95 percent of vegetation was captured during this extraction. Shadows of vegetation cannot be sepa-

rated from foliage. Percent of false alarms is very small and include local areas of unusual dark soils and shadows from few rocks, buildings and cars.

Each cluster of foliage was delivered in two formats, local (or pixel level) and integral (or model level):

- pixel level: Includes a list of each pixel and complete information about each pixel
- model level: each cluster is represented by integral parameters and a geometric primitive (ellipsoid).

File of Detailed Local Data (List of Pixels)

This ASCII-file consists of a list of pixels with parameters for each foliage object/cluster. Below is a sample of file for structure/cluster No. 291:

X	Y	Ncluster	R	G	B	Class
579	31	291	125	122	101	5
580	32	291	140	117	119	4
581	31	291	138	131	121	2
583	29	291	112	102	103	3
578	31	291	129	106	103	3

Legend:

X, Y – coordinate of each pixel of cluster (object) of vegetation (in pixels or in feet from left top corner of a tile).

Ncluster – number of cluster or structure on the tile (each tile has new numbering).

R – red component of color in this pixel

G – green component of color in this pixel

B – blue component of color in this pixel

Class of vegetation (= flag, see above).

File of Models and Integral Data

This ASCII-file consists of a list of objects (clusters of foliage) and integral parameters for each object:

Sample:

$N_{cluster}$	Line	X_c	Y_c	I_{aver}	R_{aver}	G_{aver}	B_{aver}	Black	R_{max}	Area
609	1	201	575	124	33	191	52	0.27	4.47	44

$N_{cluster}$	Line	X_{el}	Y_{el}	a	b	c	Angle (deg)
609	2	202	573	5.39	5.39	5.6	201.80

Legend:

$N_{cluster}$	–	number of clusters or structures on a tile (each tile has new numbering)
Line	–	each line presents a new set of integral parameters of cluster of vegetation and a model of cluster. Total number of lines of parameters for the object (cluster of vegetation) depends on the model and types of data used (parameters can include multispectral, radar and other data).
X_c, Y_c	-	coordinates of center of mass of cluster (all X, Y – in pixels or feet from left top corner of a tile)
I_{aver}	-	average value of intensities of all pixels of cluster
R_{aver}	–	average red component of color in this cluster
G_{aver}	–	average green component of color in this cluster
B_{aver}	–	average blue component of color in this cluster
Black	–	a part of black (I_green <10) pixels in this cluster
R_{max}	-	distance between the center of mass of cluster and most distant pixel of cluster (in pixels or feet)
Area	-	area of a cluster (in pixels or sq ft)
X_{el},Y_{el}	-	coordinates of geometrical center of cluster (average coordinates from maximal and minimal X and Y for pixels of cluster). These coordinates will be used as ellipsoid coordinates for modeling this cluster.
a	-	semi-major axis of an ellipsoid (in horizontal plane, in pixels)
b	-	semi-minor axis of an ellipsoid (in horizontal plane, in pixels)
c	–	estimation of vertical axis of ellipsoid that describes average vertical extension of cluster, in pixels (estimation from typical height of vegetation, size and shape of object, area of shadow).
Angle	-	angle of inclination of semi-major axis of ellipsoid to coordinates axis (degrees). Angle =0 for X-axis and increase to **counterclockwise** direction

The ellipsoid is considered as the zero-order approximation for foliage cluster. In many cases, such model is too simple; it does not present the complex shape of the cluster correctly. An integral model (**vint** files in data delivery) provides a general idea about the position of this cluster and its number in database. For this foliage cluster, it is possible to extract a local model with the most accurate information on pixel-level presentation (**vloc** files in data delivery) from this database.

Estimation of heights is very approximately, depends from average parameters of clusters and follows next rules (for bushes and trees):

```
For average   100<I_green<120 height = 0.1 m
For average   80<I_green<100 height = 0.5 m
For average   60<I_green<80  height = 1.0 m
For average   40<I_green<60  height = 1.5 m
```

```
For average   I_green<40  height = 2.0 m
For cluster with area > 400 and average I_green<20 height = 5.0m
For cluster with area > 1000 and average I_green<10 height = 10.0m
```

Figure 40 shows extracted clusters of vegetation (local model) and elements of integral models—ellipses contoured with each cluster of vegetation. In many cases, these ellipses overlapped.

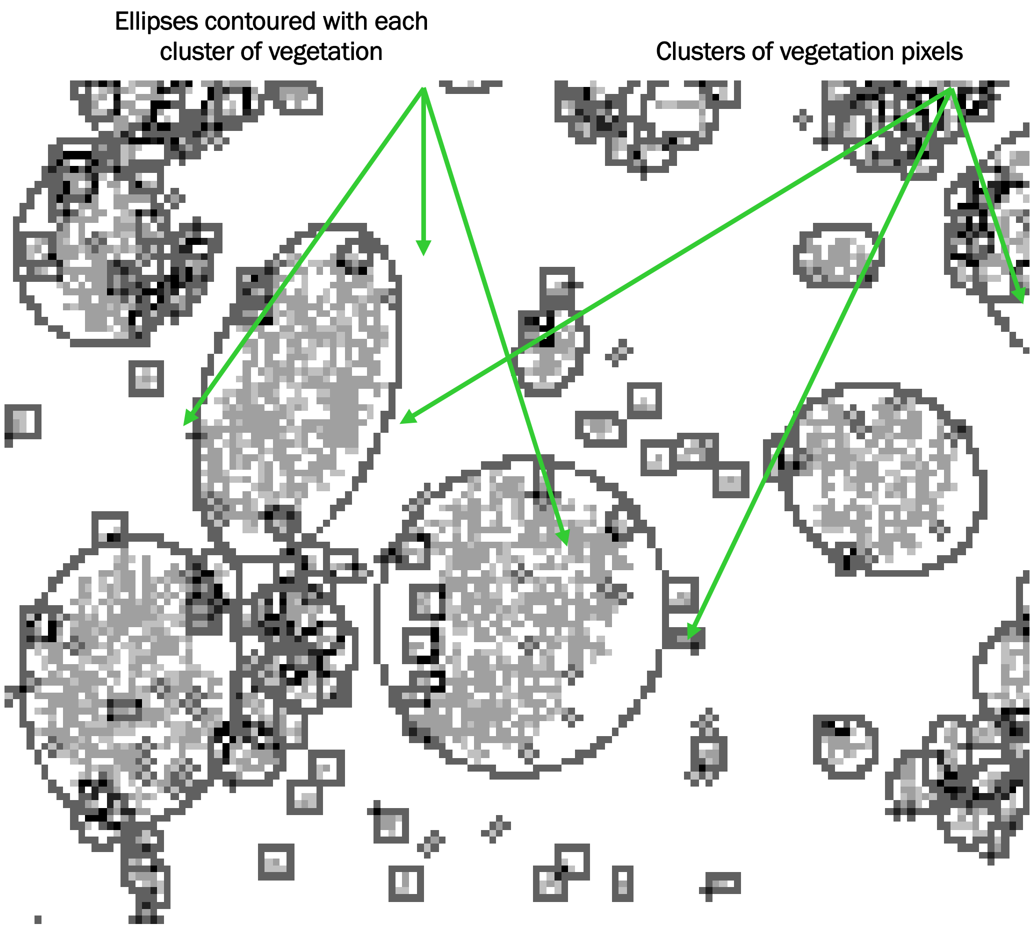


Figure 40. Extracted clusters of vegetation: dark pixels show dark vegetation, lighter pixel show other type of vegetation. Ellipses (darkest pixel) represent integral models. Black pixels illustrate where an ellipse overlapped with a pixel from another cluster; an ellipse cannot overlap with pixels of its own cluster.

Samples of Data Product

Foliage

Quality of extraction of foliage for WSMR is relatively high and can be estiamted as 90–95 percent. Number of missed objects and false alarms is small (~few percents). Figure 41 shows the typical imagery at White Sands Missiles Range. Figure 42 shows automatically extracted foliage. Some parts of very light clusters in top of imagery is missed in extracted database. Figures 43 and 44 show part of desert with rocks, that deliver some false alarms to database.

Figure 41. Imagery for white sand desert with very light vegetation (top part of imagery)

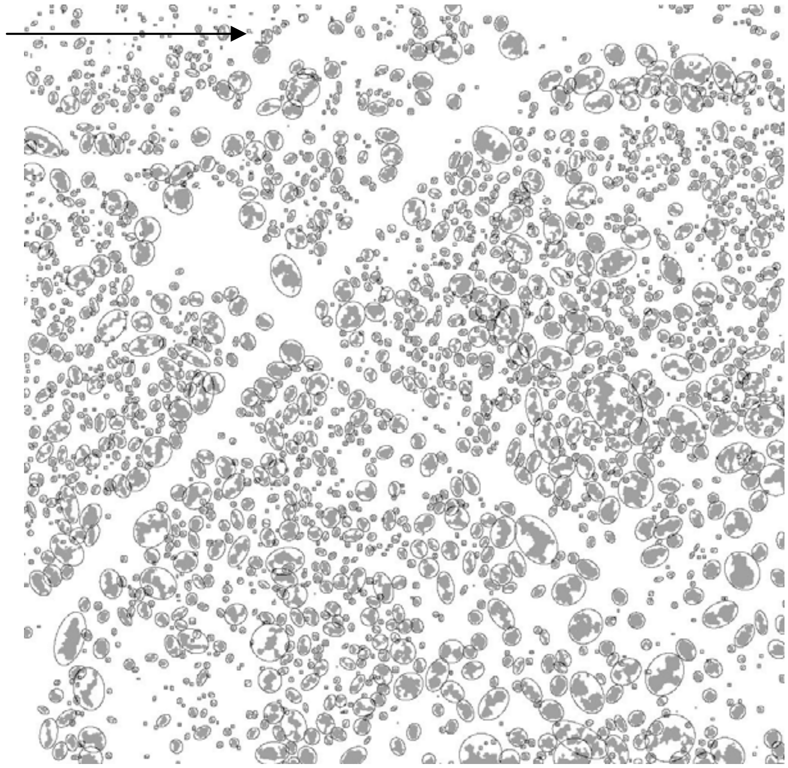


Figure 42. Extracted foliage objects for imagery from Fig. 41 with missed parts of very light vegetation clusters.

Figure 43. Grey desert with rocks and shadows of rocks.

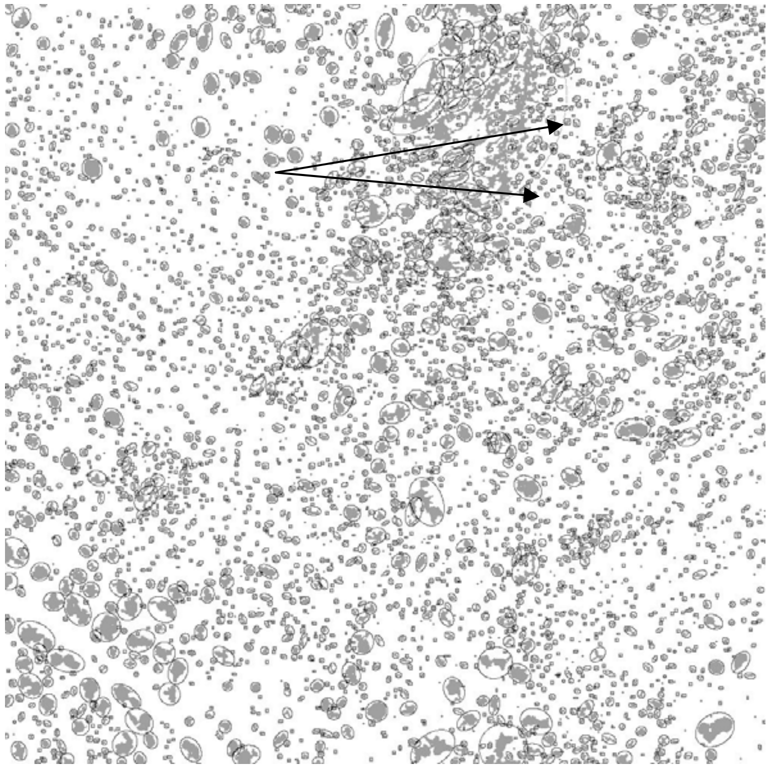


Figure 44. Extracted foliage objects for imagery from Fig. 43 with false alarms from shadows from rocks.

Figure 45 shows a good sample of large clusters of vegetation from bushes and trees. From 323 tiles of data with 30-cm resolution, 1 million and 95 thousand foliage objects were extracted. Average density of objects is approximately 3,000 per tile or 35,000–40,000 per km^2 or 1 object per 30 m^2 (around 300 pixels).

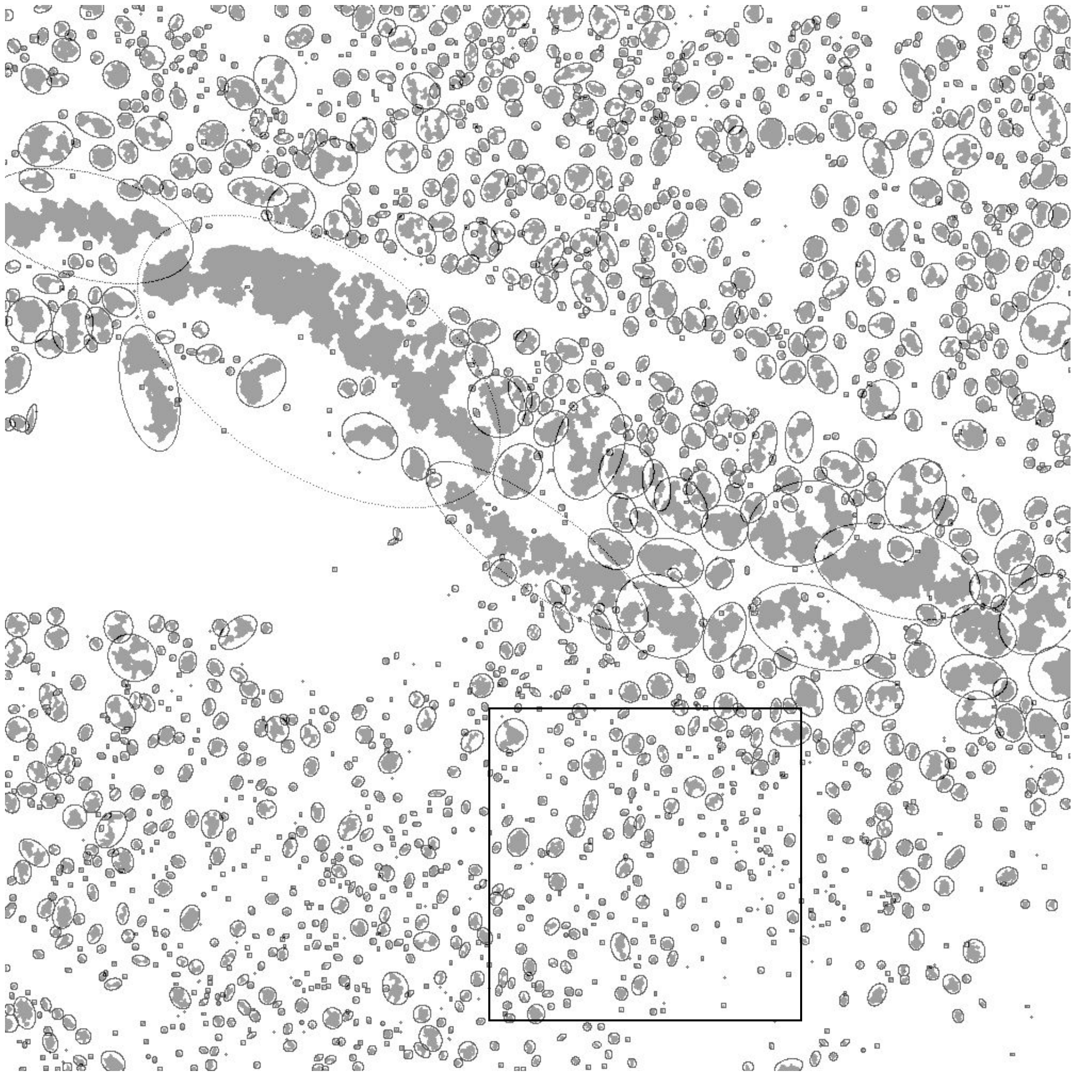


Figure 45. Extracted foliage objects for imagery from Fig. 31 with large clusters of vegetation. (Fig. 46 shows the square area 75x75.)

Figure 46 shows a comparison of vegetation and extracted models for an area of 75x75 m.

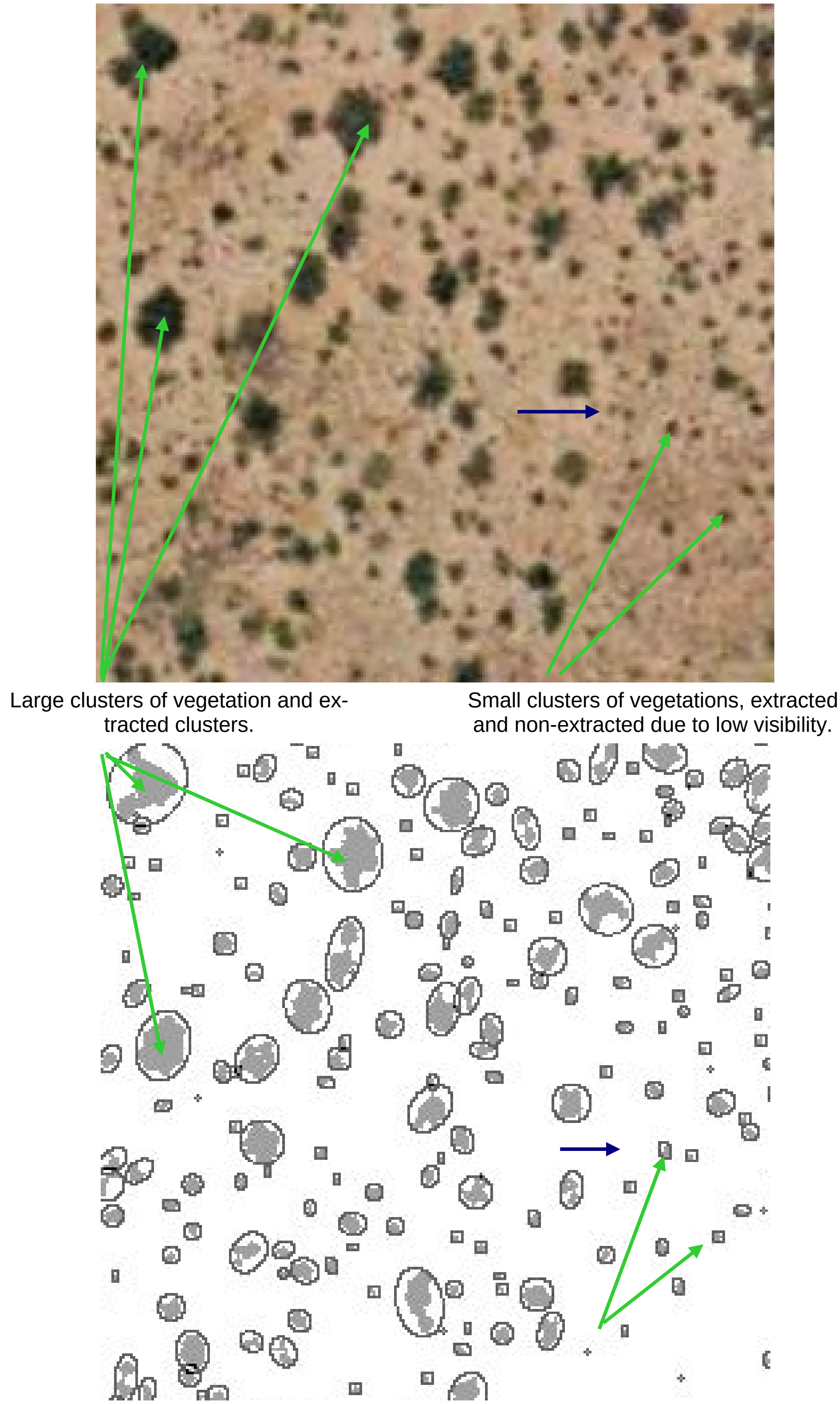


Figure 46. The comparison of vegetation and extracted models for area 75x75 m.

Roads

Road database consist all roads as polylines with minimal shape points and attributes that can be extracted from color imagery. This polyline (Figures 47 and 48) was delivered as shape file (in ArcView format).

Next, a list of attributes and characteristics of roads was provided. The attributes of roads (segment of roads) are listed as:

1. Length of road (in m)
2. Width of road (in m)
3. Type of road (Asphalt, Unpaved, Trail).

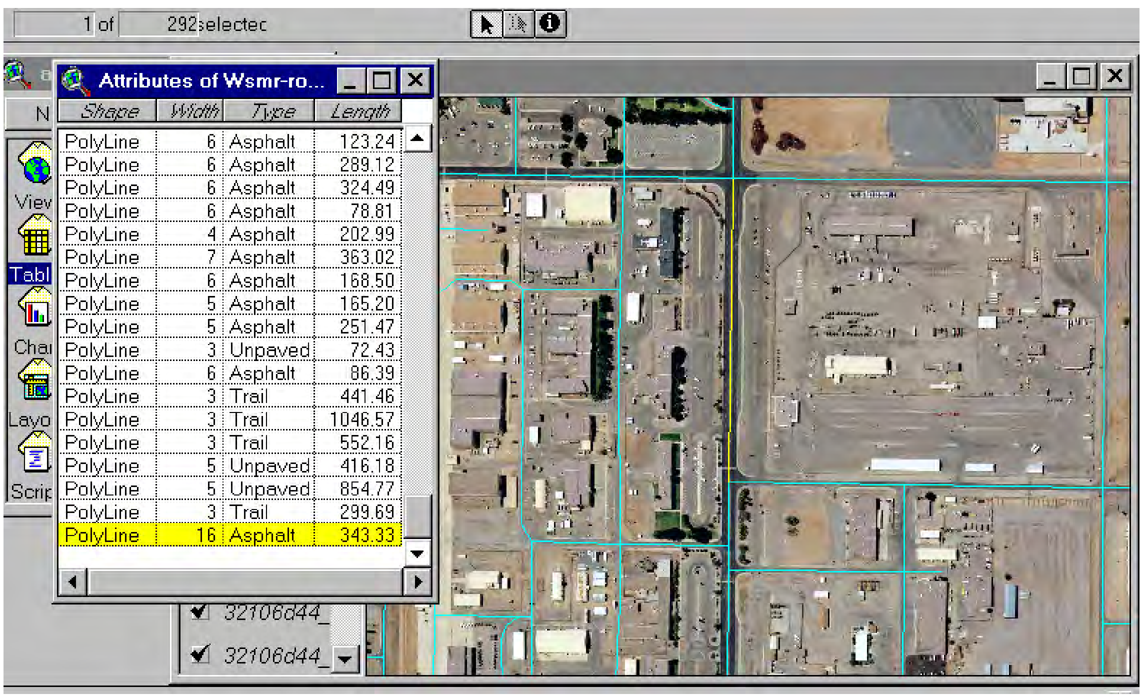


Figure 47. Imagery for urban area and according polylines for roads with attributes.

Figure 48. Whole area in 4.5 km² with extracted roads (left). Sample of unpaved roads and trails (right).

The total of 42 km of roads (292 segments), from 50 tiles of data with 30-cm resolution (4.5 km²), was extracted.

Summary and Recommendations

Summary

This project is an important step toward improving the ability to extraction of foliage objects from color imagery with accuracy of 30 cm. As result of this project, the following products were developed and delivered:

1. 353 color imagery, 300x300 m or 1000x1000 pixels. GeoTIFF. (*.tif in folder IMAGERY). Name of imagery comprising from name of original large image and coordinates of center of pixel at left top corner of imagery with accuracy 0.1 m.
2. 323 imagery with classified foliage pixels, 300x300 m or 1000x1000 pixels. GeoTIFF. (vell*.tif in folder FOLIAGE IMAGERY). Total size of processed imagery is more than 29 km².

3. 323 files with integral models of vegetation. Total number of extracted foliage obects: 1,095,000.
   (vint*.dat in folder INTEGRAL VEGETATION MODELS)
4. 323 files with local models of vegetation. Total number of extracted foliage obects: 1,095,000.
   (vloc*.dat in folder LOCAL VEGETATION MODELS)
5. Polyline of roads with attributes. ArcView format; 292 roads (segments of roads) were extracted in area of 4.5 km^2 (50 imagery, 300x300 m each). Total length of extracted roads more than 42 km.
   (ArcView files for polylines of roads in folder ROADS; imagery – in folder IMAGERY)

Recommendations

1. For the next phase of the WSMR project, color imagery can be used for foliage extraction with quality 95 percent and level of false alarms smaller than 5 percent.
2. Discussion with customers revealed the necessity to improve an algorithm for the extraction and determination of foliage objects, especially in urban area with green lawns.
3. Improvement of the vegetation models can be performed after:
   a. Improving the criteria for extraction of foliage against backgrounds with variability of brightness and spectral parameters.
   b. Improving the estimation for heights of foliage using field tests and control measurements of heights of different types of vegetation.

4 Verification and Validation Assessment of Interferometric Synthetic Aperture Radar AFE

Executive Summary

The overall objective of this part of the study was to perform a Verification and Validation (V&V) assessment of IFSAR Automatic Feature Extraction (AFE) on the currently available IFSAR Digital Surface Model (DSM) and Digital Terrain Model (DTM) for the Yuma Wash area. As part of this work, the Government furnished the CCS team with IFSAR data and control points for the Yuma Wash Area to partially validate the data. Four quantitative parameters were tested and used to analyze the quality and fidelity of the IFSAR, AFE and DTM:

1. RMSE, which must be minimal
2. Bias, which ideally should be minimal and close to zero
3. Maximal percent of extracted artificial objects and trees
4. Maximal percent of maintained details of relief.

The final results show that using this V&V method and AFE algorithms can make a significant improvement to the Intermap DTM (Table 5). In all cases, the data improved statistically, and the overall data integrity of the true surface elevations was also maintained. Note that the relative accuracy of the Intermap DSM is of good quality as a standalone product; the absolute accuracy of this dataset was in disagreement only with the control points. This type of error is not believed to be representative of all IFSAR data.

Table 5. Assessment of AFE and DTM from Intermap and CCS.

Final Statistics Comparing Original Data with CCS Processed Data		
Criteria of AFE	Intermap AFE and DTM	CCS AFE and DTM
RMSE	130 cm	65 cm
Bias	+104 cm	+ 3 cm
Percent of extracted objects	Estimated <90% (no quantitative measure)	Estimated >90% (no quantitative measure)
Percent of maintained details of relief	Estimated <50% (no quantitative measure)	Estimated 90–100% (no quantitative measure)

Introduction

TEC-OFC contracted Computational Consulting Services, LLC (CCS) in March 2006 to perform a V&V assessment of IFSAR AFE on the current available IFSAR DSM and DTM for the Yuma Wash area. The following issues were analyzed to assess the fidelity and quality of the DTM and AFE:

1. Overall accuracy by measuring the amount of error between a set of ground-truth control points, and the DSM/DTM using the Root Mean Square Error methodology. Using the RMSE not only analyzes the accuracy, but also identifies potential shifts or biases in the data. – **RMSE and Control Points (1).**
2. Percentage of Extracted Artificial Objects and Foliage – for performing foliage/vegetation removal and building/structure filtering. – **Percent of Extracted Artificial Objects and Foliage (2).**
3. Percentage of Extracted Details of Relief – for maintaining accurate terrain profile – **Percent of Maintained Details of Relief (3).**

RMSE and Control Points (1)

In December 2004, TRAC-WSMR conducted a field survey of the Yuma Wash area in which GPS surveys were performed. This survey consisted of collecting ground control points along five lines for a total of 2481 control points (see CCS report *High Resolution Terrain Modeling Using Interferometric Synthetic Aperture Radar (IFSAR) Data* regarding assessment of accuracy). Since one of the lines showed a bias and did not fit internally against the other lines in the survey, only four of the five lines were used. Even after removing one of the lines, there were still 2186 control points, which are easily sufficient to assess the quality of the data. By comparing the control points with the IFSAR DSM and DTM, a Root Mean Square Error(z) is computed for each model. The statistics from Table 6 indicate a systematic degradation of RMSE in the DSM vs. the DTM. Since the control points were not in heavily vegetated areas, the DTM should be relatively close to that of the DSM with the same minimum and maximum errors, and this clearly is not the case.

Table 6. RMSE and maximal errors between control points and the DSM/DTM.

RMSE and Maximal Errors Based on Control Points					
Control Point Line(s)	# of Pts	Surface Type	RMSE (cm)	Min Error (cm)	Max Error (cm)
Line 728	684	DSM	118	-263	355
		DTM	125	-251	401
Line 729	736	DSM	124	-114	317
		DTM	137	-359	323
Line 3665	417	DSM	131	-103	383
		DTM	132	-125	430
Line 3667	349	DSM	117	-154	357
		DTM	118	-271	386
All 4 Lines	2186	DSM	123	-263	383
		DTM	130	-359	403

By assessing the RMSE, a second parameter is also studied, i.e., if a potential shift or bias exists within the data. This analysis summarizes that a bias does exist between the Intermap DSM and the control points on the magnitude of +103.8 cm (Figure 49). Taking this shift into account, this process lowered the DTM approximately 1 m such that the results significantly improve to ~ 0.03 m (Figure 50).

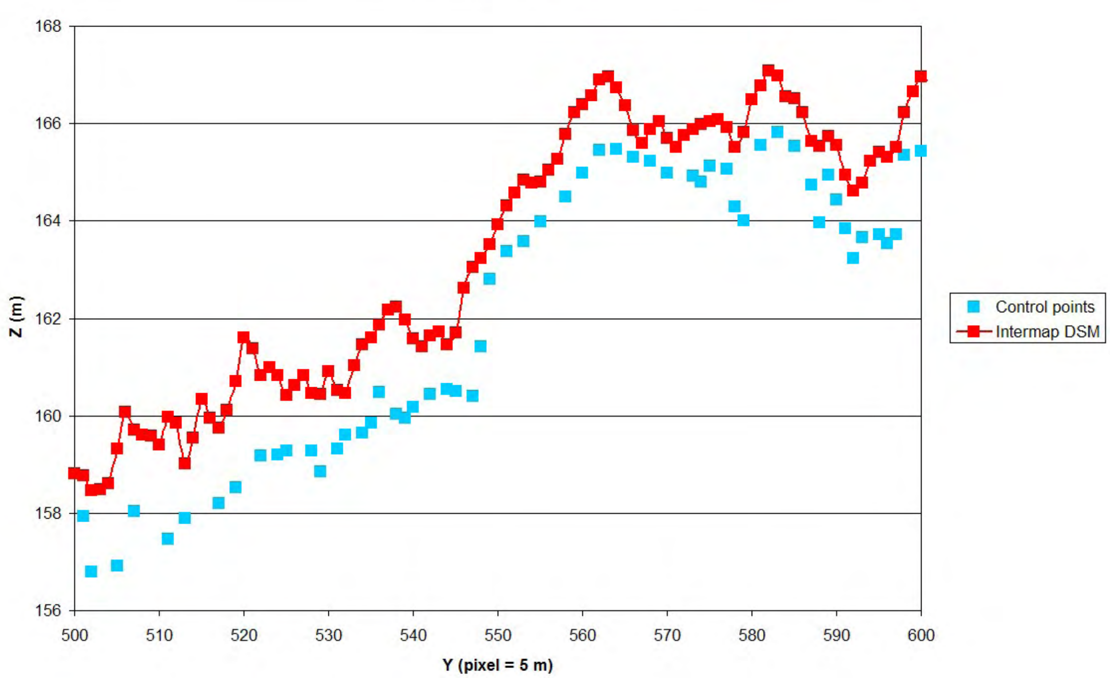


Figure 49. Intermap DSM and control points for 500 m (= 100 pixels) of line No. 729. Blue squares are control points, red squares are DSM. Bias of Intermap DSM from 2186 control points is +103.8 cm.

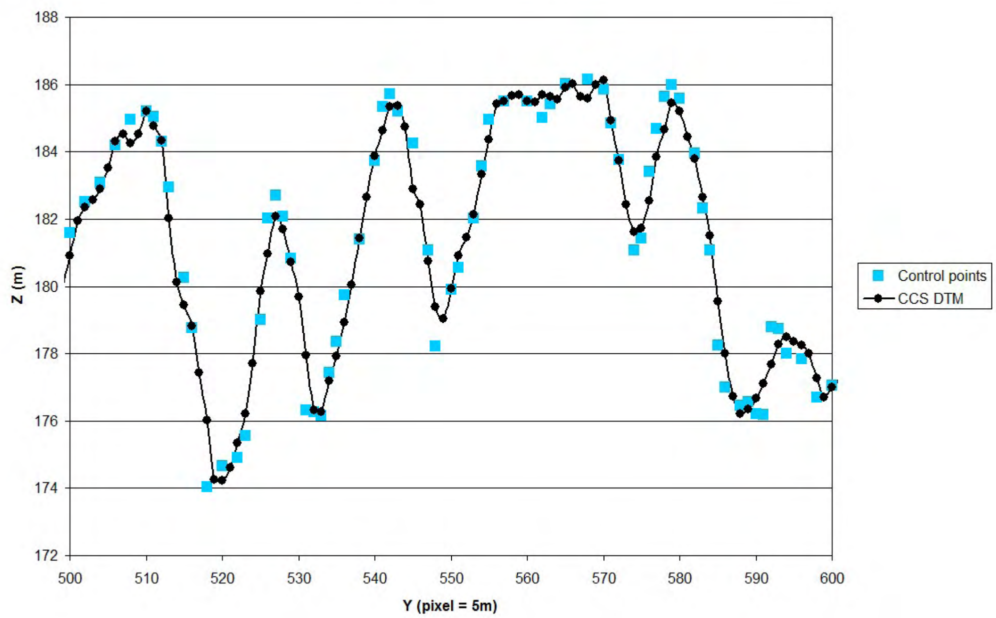


Figure 50. CCS DTM and control points for 500 m (= 100 pixels) of line No. 728. Blue squares are control points, black circles are DTM. Bias of CCS DTM from 2186 control points is +3.1 cm.

For analyzing the accuracy and fidelity of the IFSAR DTM, two quantitative parameters are measured:

- RMSE, which must be minimal
- bias, which must be minimal or close to zero for LIDAR and IFSAR data.

Calculation of the RMSE, bias, maximal, and minimal errors done in this V&V software (Figure 51), which also calculates the difference between DSM and DTM, and which will be used in next steps of this V&V assessment of IFSAR AFE.

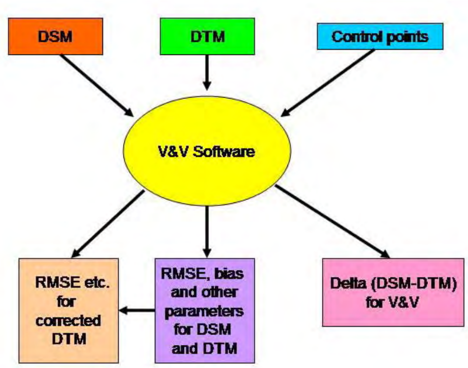


Figure 51. Algorithm process flow of V&V of IFSAR data for this project.

Percent of Extracted Artificial Objects and Foliage (2)

Control points, calculation of the RMSE and the bias, are very important tools for improving IFSAR topographic data products, but these criteria are not sufficient to guarantee a quality DTM. The RMSE can measure the quantitative aspect of the data, but it does not measure the qualitative aspect, or how well the product conforms to a true bare-earth product unless some of the control points are in vegetated areas. This results in vegetation having a minimal impact on the RMSE (for this dataset). One component of the V&V assessment of AFE includes an estimation of the percent of extracted artificial objects and foliage from the DSM to create the DTM. To illustrate this process; data from a previous CCS project for Dewberry/FEMA is presented. This data consists of a LIDAR DSM with 16 ft (4.8 m) pixel resolution (which is similar in size to the IFSAR DSM) for Mecklenburg County, NC. Figure 52 shows the LIDAR DSM with numerous buildings. Figure 53 shows the same area, but using the CCS processing method to create the DTM that is significantly free from artifacts except for a few minor areas.

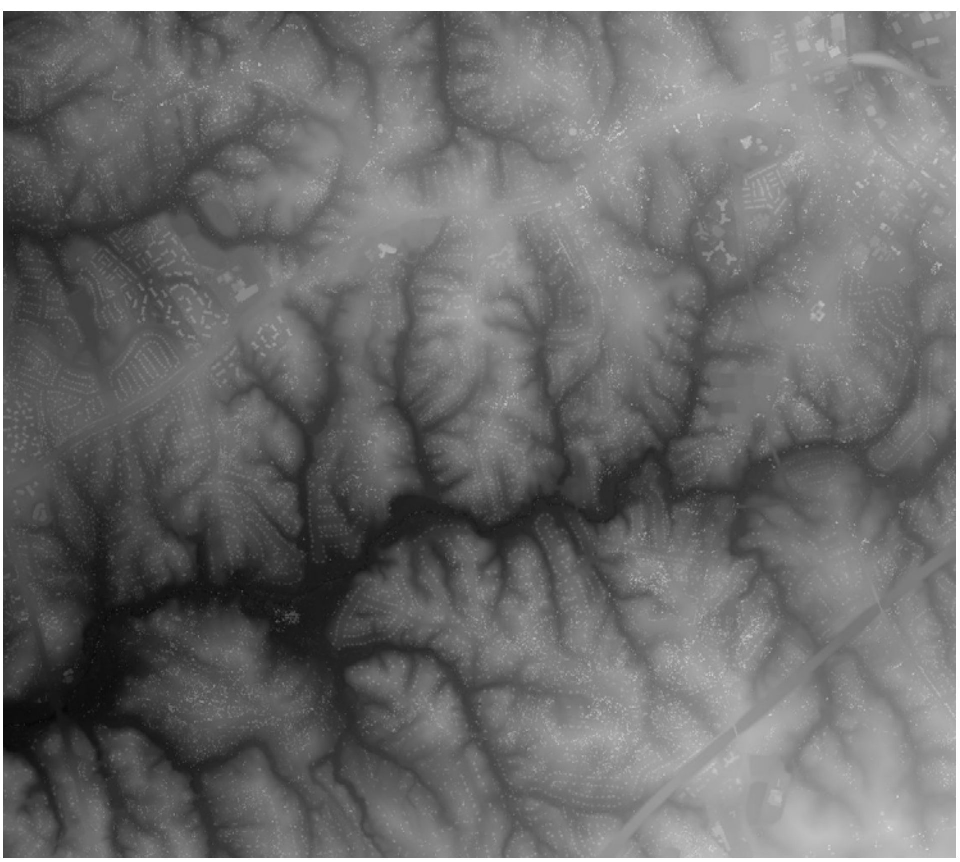


Figure 52. Sample from CCS project (2004) for Dewberry/FEMA. LIDAR DSM with 16 ft resolution for Mecklenburg County, NC. Size of tile is 18,000 x 16, 000 ft.

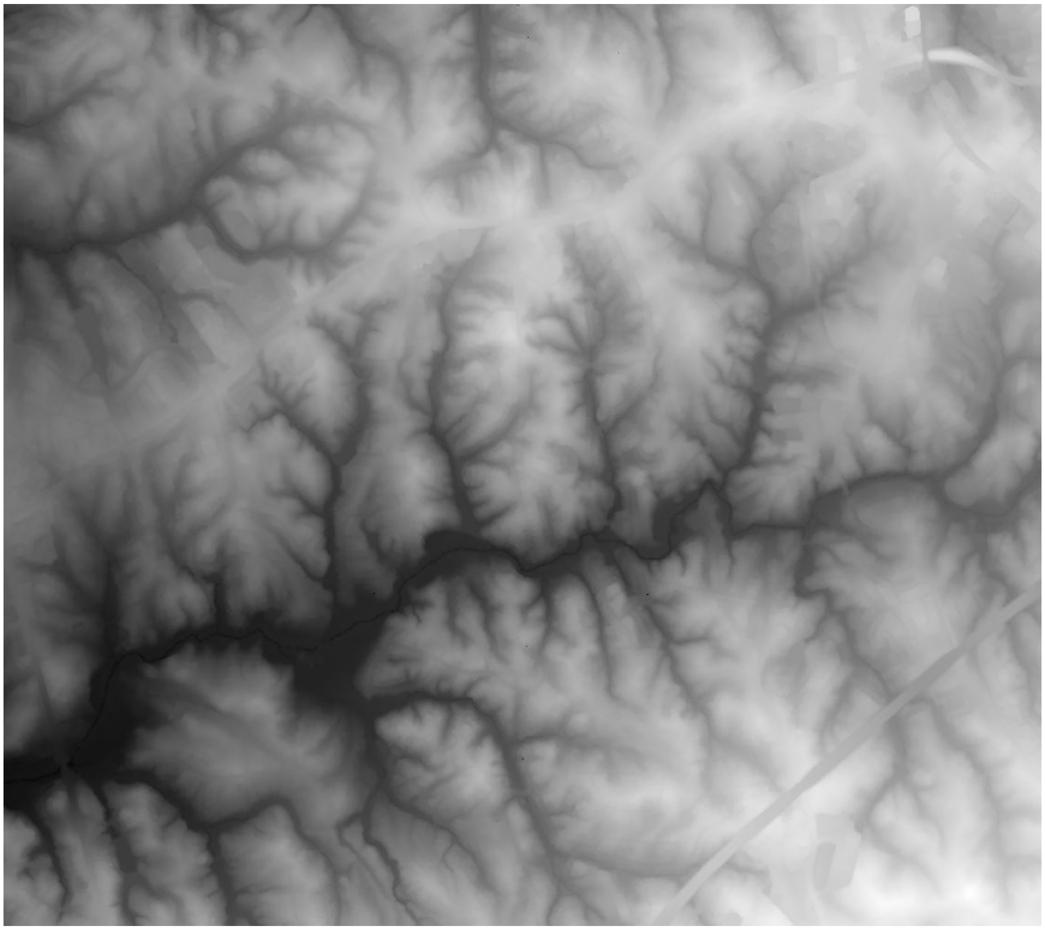


Figure 53. Sample from CCS project (2004) for Dewberry/FEMA. CCS DTM from LIDAR DSM shows one or two parts of non-extracted buildings (for example, on top right corner).

To verify these results, the planimetrics of the building footprints were compared to the extracted features of this data set. In total, the planimetrics identified 13,500 structures, which included houses, garages, and some sheds. The AFE process was able to identify 12,804 objects, which results in a success rate of 94.8 percent.

This analysis of DTM and the estimation of percent of extracted objects are important criteria for determining the quality of AFE. Therefore, a third criterion of good quality of DTM is:

- Maximal percent (90–100 percent) of extracted artificial objects and trees.

Note that in areas of steep terrain such as mountains, the estimation of the quality of the AFE through visual interpretation is difficult and additional refinement of parameters and processing may be required.

Percent of Maintained Details of Relief (3)

The above three criteria are necessary, but not always sufficient, for a good quality DTM. In many cases, aggressive filtering not only removes legitimate artifacts such as buildings and trees, but also can remove details of natural or artificial relief: dams, hills, and parts of slope surfaces. For example, areas of steep terrain with vegetation along stream channels can sometimes be aggressively filtered resulting in a wider channel, or areas of varying slope can sometimes be over smoothed. A balance must be found that maintains the integrity of the data yet yielding a representative terrain surface. Using only the RMSE cannot test for this, therefore the last criterion for a high quality AFE and DTM is:

- maximal percent of maintained details of relief.

To analyze the level of maintained details, a comparison is made between a DSM and DTM by measuring the difference of the extracted features. For example, Figure 54 shows the difference between a DSM and DTM for Mecklenburg County, NC. Only buildings, bridges and trees are presented, and no erroneous terrain values were extracted (legitimate land features). Here a good balance was found of removing legitimate artifacts while maintaining a true bare-earth terrain model.

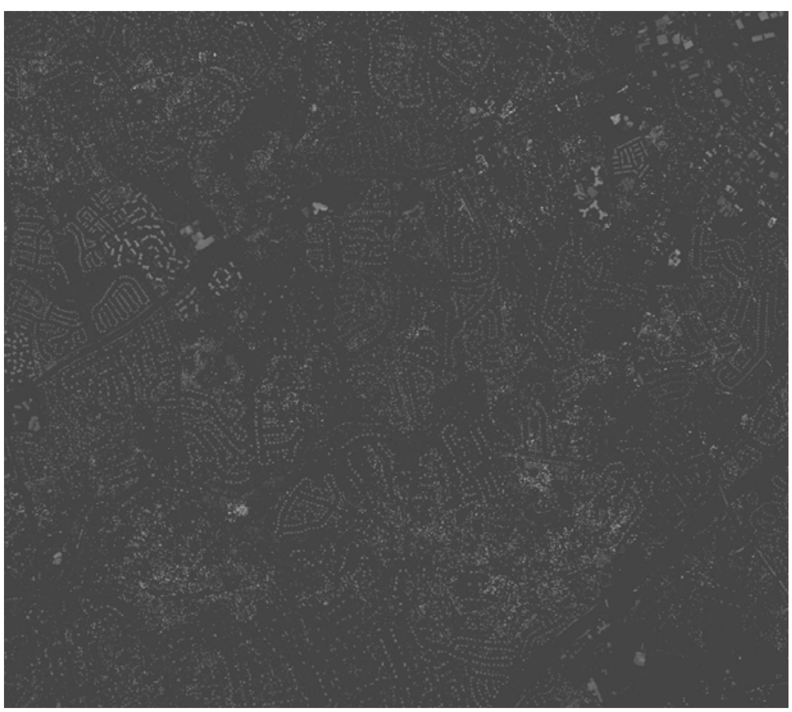


Figure 54. Sample from CCS project (2004) for Dewberry/FEMA. Difference DSM-DTM show 12,804 extracted buildings and a numerous vegetated areas. Details of natural relief are not presented.

Another example of analyzing extracted features and maintained details of relief using different processing methods can be seen with the IFSAR DSM and DTM from this project. Figure 55 shows the DSM for Yuma Wash area. Figures 56 and 57 shows the differences between the DSM and Intermap derived DTM in both positive and negative differences. Both figures illustrate a high level of details of relief (hills and hollows), which were smoothed or filled by the Intermap AFE.

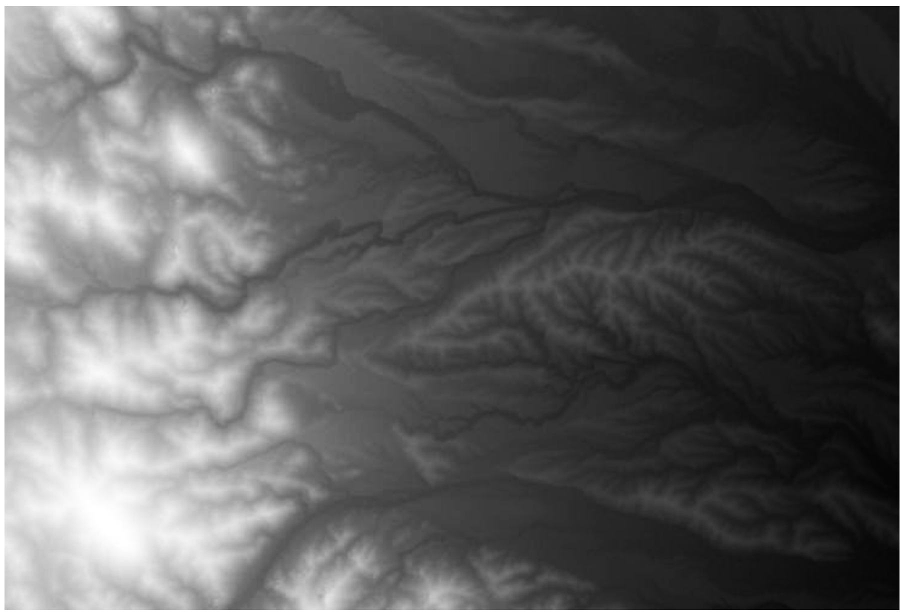


Figure 55. Intermap DSM, Yuma Wash area. There is minimal artificial objects and heavy foliage. The DTM is very similar to DSM.

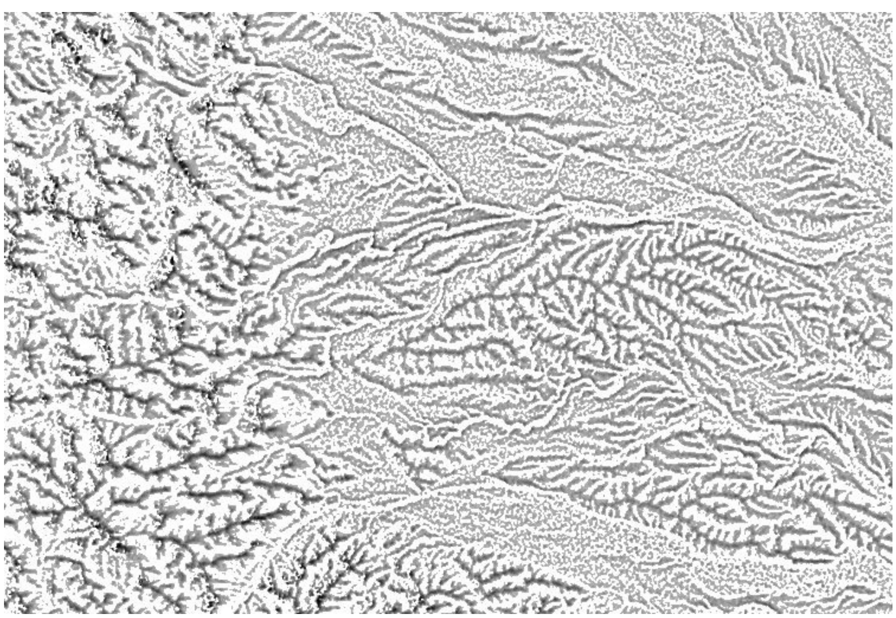


Figure 56. Illustrates the positive differences between the Intermap DSM and DTM. Note the level of smoothed relief particularly the tops of hills. This data contained minimal artificial objects.

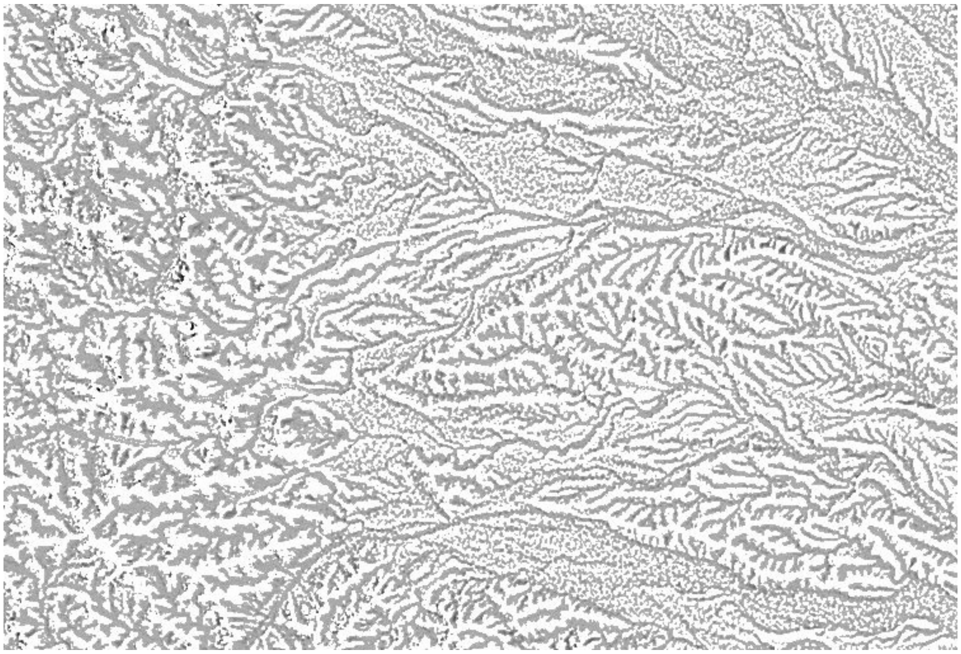


Figure 57. Illustrates the negative differences between the Intermap DSM and DTM. A lot of details of natural relief were filled, including all hollows.

Figure 58 shows CCS DTM and a fine balance between removing foliage and still retaining the finite features of the terrain. To obtain excellent AFE methods, the processing algorithm must be flexible and robust to extract legitimate features from gentle terrain such as the desert area for Yuma Wash, to very heavily forested geographic areas.

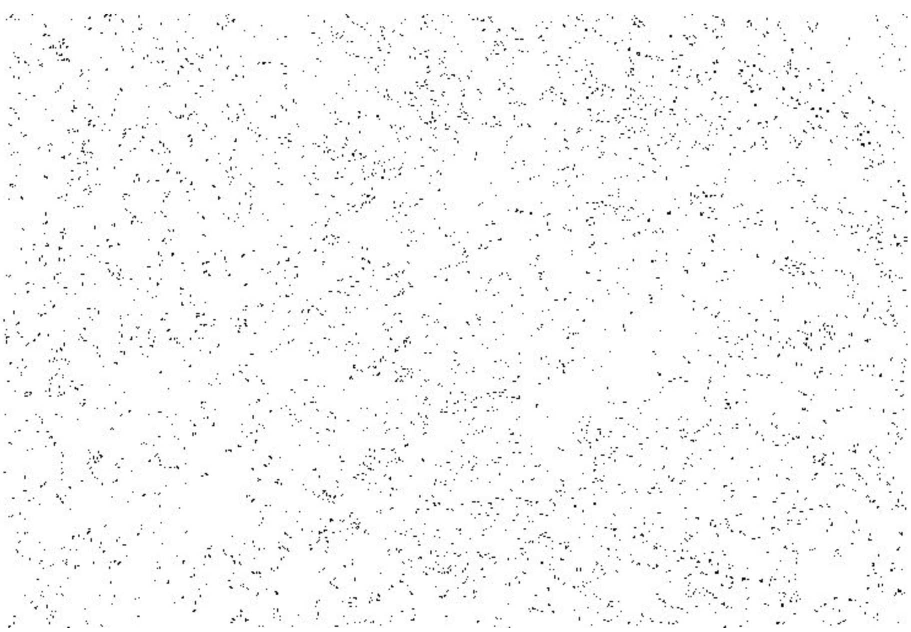


Figure 58. Difference Intermap DSM (corrected after V&V) – CCS DTM. This difference has only positive value. Only clusters of desert foliage were extracted and, may be, single large rocks. Details of natural relief are not presented.

Using this criteria and methods for V&V of IFSAR AFE has significantly improved the accuracy of DTM from the original IFSAR data (Table 7).

Conclusion

To analyze quality and fidelity of IFSAR, AFE, and DTM, requires the use of quantitative parameters or criteria of quality AFE:

1. RMSE, which must be minimal
2. Bias, which ideally should be minimal and close to zero
3. Maximal percent of extracted artificial objects and trees
4. Maximal percent of maintained details of relief.

These AFE algorithms and V&V method can together significantly improve the Intermap DTM and provide a DTM of higher quality. The RMSE improved from over 100 cm to 65 cm. The bias was also identified and removed resulting in the improved RMSE. In all cases, the data improves statistically, and the overall data integrity of the true surface elevations is also maintained by ensuring legitimate features are extracted and the details of the relief are maintained.

Table 7. Assessment of AFE and DTM from Intermap and CCS.

Final Statistics Comparing Original Data with CCS Processed Data		
Criteria of AFE	Intermap AFE and DTM	CCS AFE and DTM
RMSE	130 cm	65 cm
Bias	+104 cm	+ 3 cm
Percent of extracted objects	Estimated <90% (no quantitative measure)	Estimated >90% ? (no quantitative measure)
Percent of maintained details of relief	Estimated <50% (no quantitative measure)	Estimated 90–100%

5 Light Detection and Ranging and Imagery Data Collection Requirements To Support the Generation of Synthetic Scenes and Digital Terrain Models

Background

The Yuma Proving Ground (YPG) is supports testing of the Army's Future Combat System (FCS) during a series of Distributed Test Events (DTEs). FCS units of action operate in a defined Common Operating Area (COA) within the training range or proving ground. To support the test and evaluation of the FCS systems, the Government requires very detailed and accurate geospatial information describing the terrain, surface features, and other objects located within the COA.

The General Dynamics Armament and Technical Products (GDATP) team that includes the Greenwich Institute for Science and Technology (GIST) and Dewberry, LLC (Dewberry) are supporting YPG and the Army Test and Evaluation Command (ATEC) by developing digital terrain and surface feature models using Light Detection and Ranging data (LIDAR) (see Section 0) that was collected from an airborne platform. During the processing of LIDAR data collected by the Government for a 100-km^2 area at White Sands Missile Range, the team identified deficiencies that affected the overall quality of the delivered data models.

This document puts forth a recommended set of requirements for collecting remote sensing data to produce geospatial data products that meet the overall program goals. Compliance with these requirements for future data collections will greatly improve the overall quality and accuracy of the geospatial data products derived. Specific requirements address pre-collection activities such as number and placement of ground control points, aircraft flight pattern plans, and instrument calibration; raw data collection and pre-processing including point density, collection controls, information content; and independent validation and verification.

Included in this requirements document is a reference guide for specifying data collection requirements and the effect of these requirements on cost, data accuracy and quality, and data model resolution and fidelity.

LIDAR Data Collection

Airborne LIDAR is an optical-based measurement system where a laser is mounted to a rotary- or fixed-wing aircraft, that acquires x, y, and z coordinates of terrain and terrain features, both manmade and naturally occurring. LIDAR systems consist of an airborne Global Positioning System (GPS) with attendant GPS base station(s), Inertial Measuring Unit (IMU), and light-emitting scanning laser. The system measures ranges from the scanning laser to terrain surfaces within a scan width beneath the aircraft. The time it takes for the emitted light (LIDAR return) to reach the earth's surface and reflect back to the onboard LIDAR detector is measured to determine the range to ground. Scan widths will vary, depending on mission purpose, weather conditions, desired point density and spacing, and other factors. The other two components of LIDAR systems are the airborne GPS, which ascertains the in-flight three-dimensional position of the sensor, and the IMU, which delivers precise information about the attitude of the sensor.

Equipment

First and Last Returns

The LIDAR equipment shall measure both first and last returns for each pulse. Additional pulses are desired, but not mandatory.

Intensity Measurements

The LIDAR equipment shall measure the intensity of each return.

Calibration

LIDAR systems shall be fully calibrated to the manufacturers' specifications after installation or reinstallation and follow a fully approved calibration testing methodology.

LIDAR systems shall be calibrated on-site prior to each acquisition mission to ensure correct elevations are measured when flown from all cardinal directions and to obtain the correct bore-sighting (horizontal positioning of laser pulses on the ground). Post mission validation checks are also necessary prior to moving to the next project area to ensure complete coverage with acceptable data.

Artifacts

Artifacts are regions of anomalous elevations or oscillations and ripples within the Digital Elevation Model (DEM) data resulting from systematic errors or environmental conditions. They may result from malfunctioning sensors, poorly calibrated instrumentation, adverse atmospheric conditions, or processing errors. Whereas positive (above ground) local artifacts can result from errors in post-processing to remove buildings and vegetation, negative (below ground) local artifacts result from errors in sensor calibration.

Artifacts in the LIDAR data must be kept to a minimum. Artificial underground reflections deeper than 30 cm must be <1 per 10,000,000 shots for normal geographic areas and <1 per 1,000,000 shots for an urban area or feature transition areas (e.g., body of water to land transition). Artificial positive reflections higher than 30 cm must be <1/1,000,000. Figures 59 and 60 show samples of excessive local artifacts.

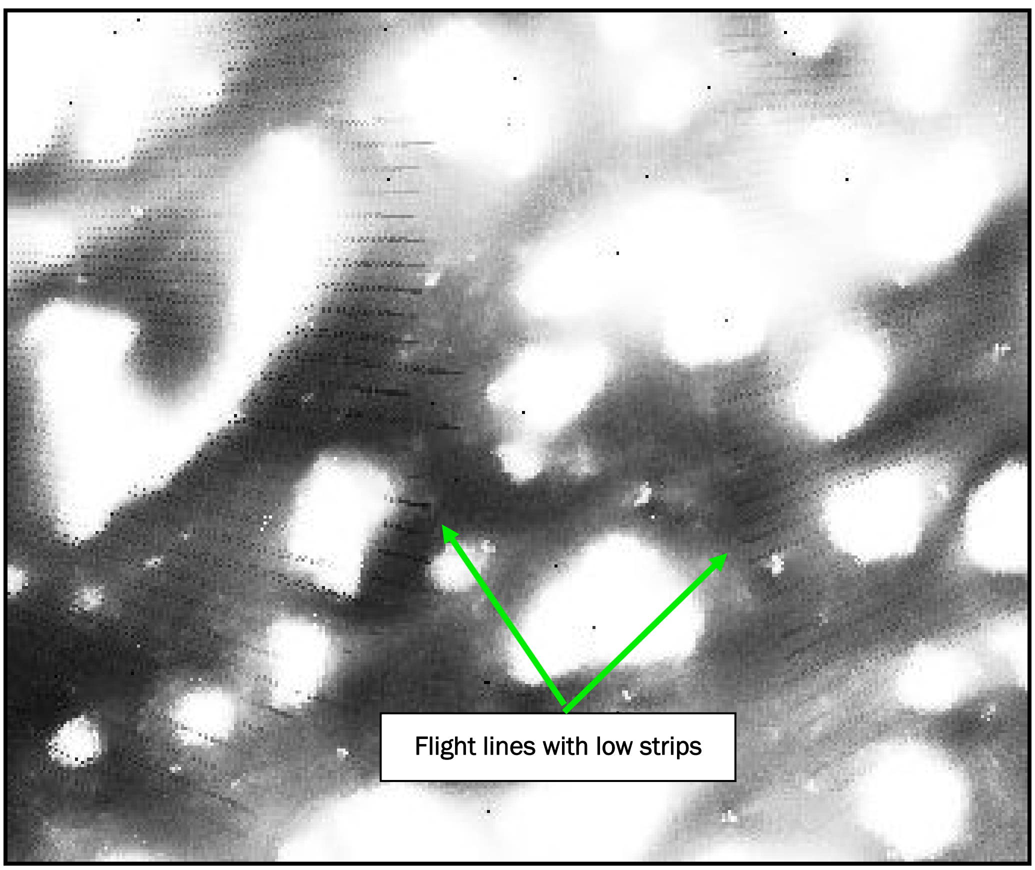


Figure 59. Flight lines with strip-like artifacts, underground and above-ground "corn rows."

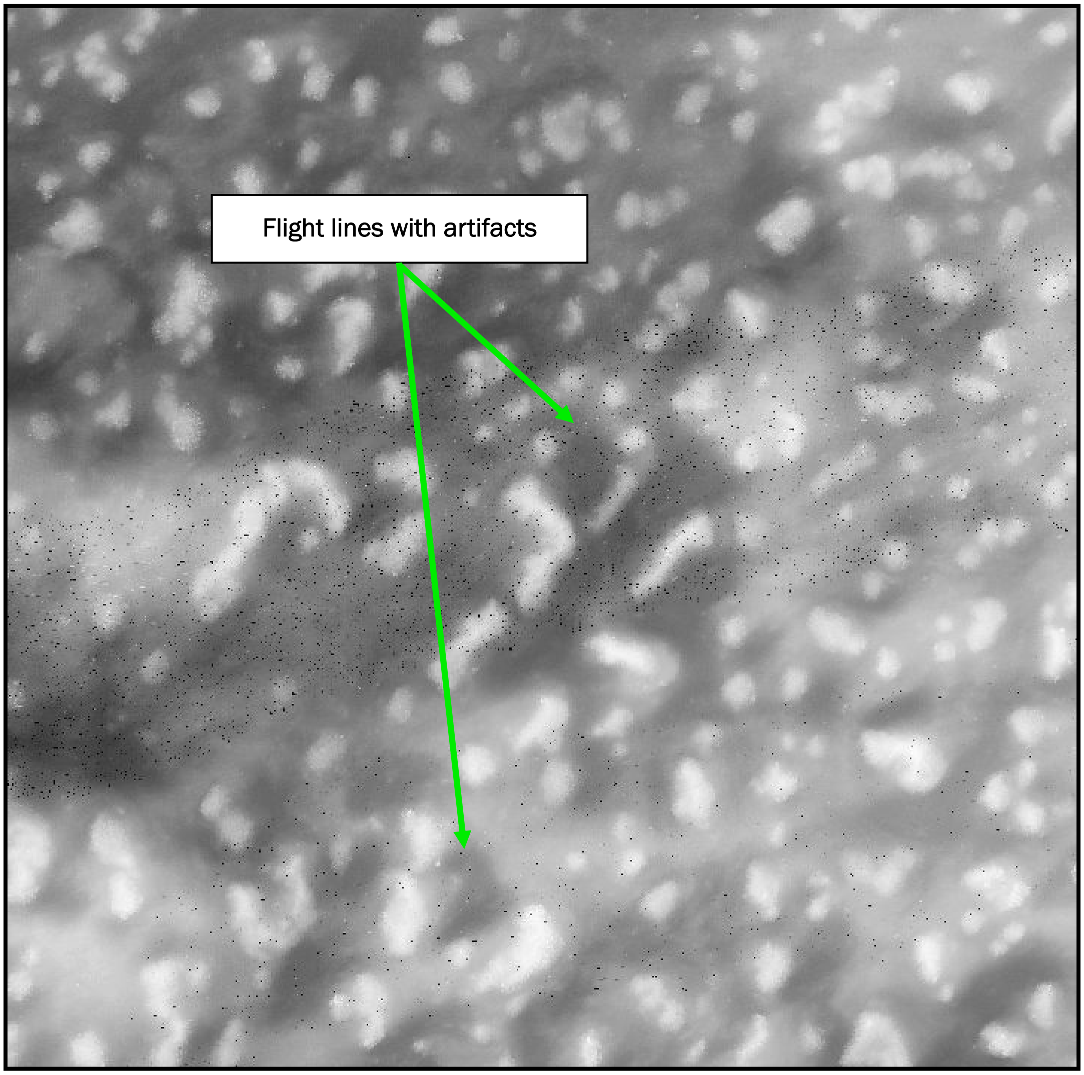


Figure 60. Flight lines with excessive point-like negative artifacts, underground reflections.

Vertical Accuracy

LIDAR systems operated from commercial fixed-wing aircraft typically acquire data from a flight altitude of approximately 1000 m above mean terrain with nominal point spacing of 1.5 m, less than one LIDAR pulse for each 1-m pixel of associated imagery. The following vertical accuracy requirements apply to LIDAR data collected using fixed-wing aircraft:

- absolute vertical accuracy = 15 cm Root Mean Square Error (RMSE) = 29.4 cm (11.6 in.) vertical accuracy at the 95 percent confidence level.
- relative vertical accuracy between LIDAR points in the same flight line is 5–7 cm at the 95 percent confidence level.
- relative vertical accuracy between LIDAR points in overlapping flight lines is 7–10 cm at the 95 percent confidence level.

LIDAR systems operated from helicopters typically acquire data from lower altitudes with sub-meter point spacing, i.e., more than one LIDAR pulse for each 1-m pixel of associated imagery. The following vertical accuracy requirements apply to LIDAR data collected using rotary-wing aircraft:

- absolute vertical accuracy = 10 cm RMSE = 19.6 cm (7.7 in.) vertical accuracy at the 95 percent confidence level
- relative vertical accuracy between LIDAR points in the same flight line is 3–5 cm at the 95 percent confidence level
- relative vertical accuracy between LIDAR points in overlapping flight lines is 5–7 cm at the 95 percent confidence level.

Figure 61 shows an unacceptable elevation "jump" along the boundary between adjoining flight lines.

Horizontal Accuracy

The required horizontal accuracy must be 1.0 m or better at the 95 percent confidence level. A simple method for estimating horizontal accuracy is to compare the horizontal shifts of sharp details (building/bridge edges) on overlapping flight lines. The maximum shift in horizontal position should be 1 pixel only. Differences of 2–3 pixels show poor horizontal consistency, which is a good indicator of poor horizontal accuracy.

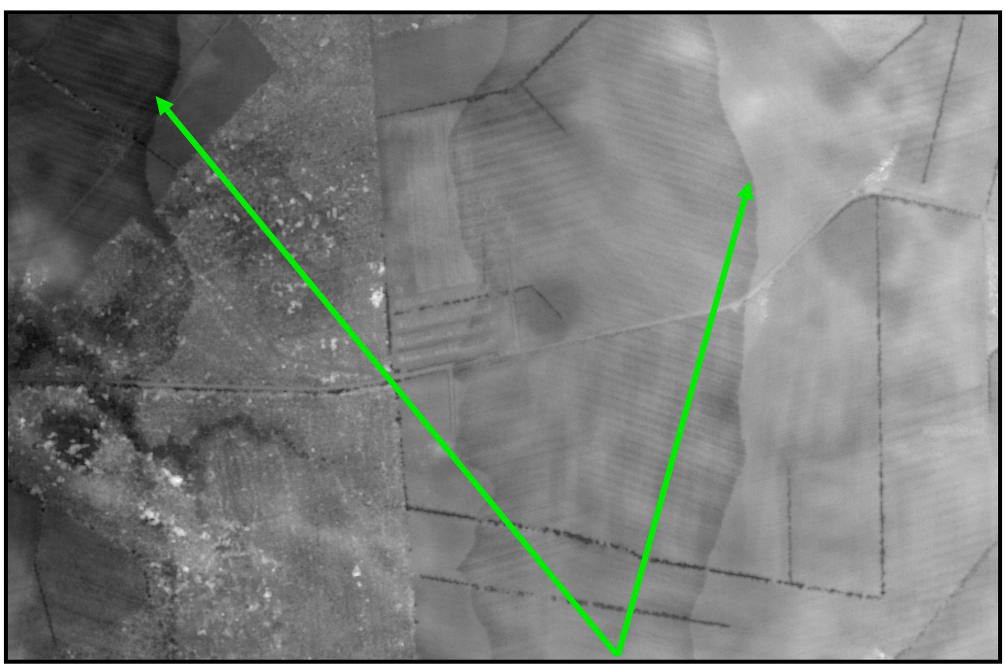


Figure 61. Sharp jumps show poor calibration of heights between adjoining flight lines.

LIDAR Sensitivity

LIDAR systems should possess high sensitivity and good penetration of vegetation. Simple criteria for estimation of sensitivity of LIDAR equipment:

- Dark roofs and asphalt roads must be visible on LIDAR data. Figure 62 shows sample data from LIDAR equipment with low sensitivity.
- For some applications, electric wires must be visible in LIDAR first returns (Figure 63).
- Flight parameters must be optimized to maximize the penetration of foliage for generation of bare-earth elevations. Figure 64 shows the discrepancies caused when the LIDAR dataset is part leaf-off and part leaf-on, with poor penetration.

LIDAR Pulse Repetition Rate

The typical frequency of LIDAR measurements for the best commercial LIDAR systems is 33,000–50,000 laser pulses per second with the newer systems measuring 100,000 laser pulses per second.

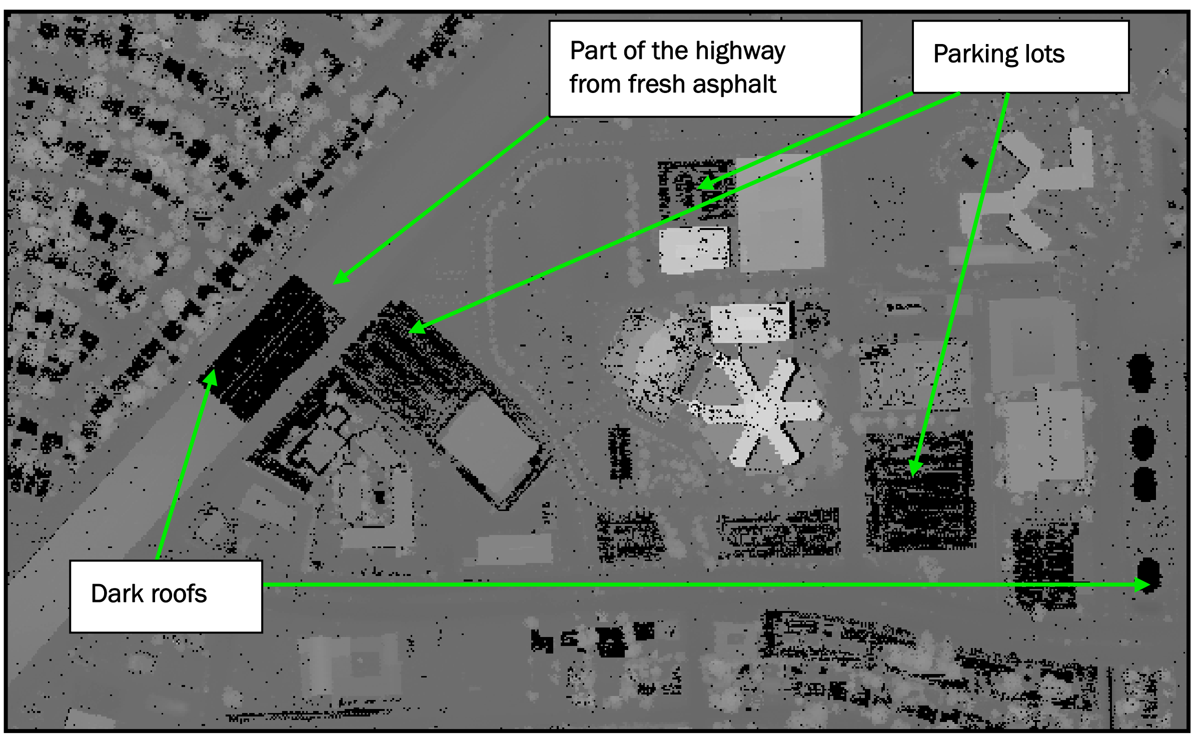
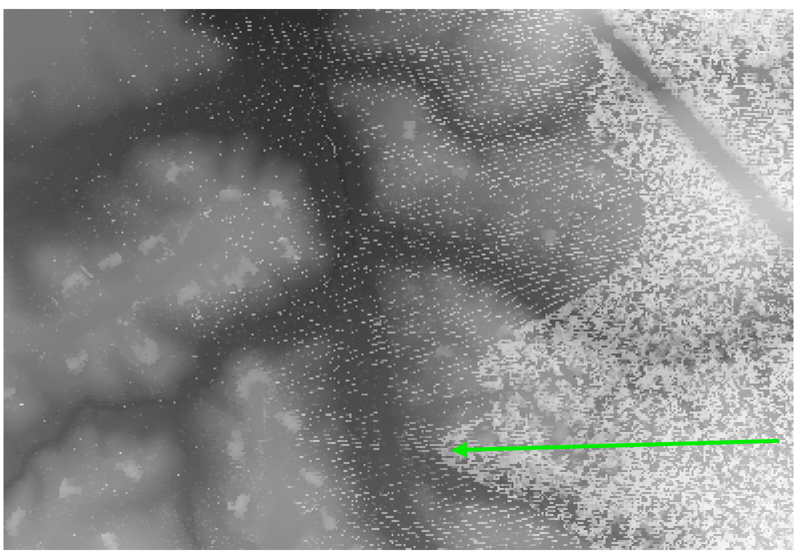


Figure 62. LIDAR last returns. All dark areas are areas without LIDAR data. These dark areas include parking lots, part of a highway with fresh asphalt, and homes with dark roofs.

Figure 63. Most electric wires should be visible on first returns from LIDAR equipment with high sensitivity.



This image presents minimal heights of LIDAR last returns. The pixel size is 2 m. Two databases of LIDAR data were combined in this tile, sized 790x540 m. LIDAR data for the left side was collected in early spring during leaf-off state; data for the right side was collected later when trees were in leaf-on state. The middle part of tile shows both databases. The boundary of forested area reflects not the actual edge of forest, but the boundary of the data.

Figure 64. Discrepancies between leaf-on and leaf-off data.

Pre-Collection Activities

For a LIDAR survey, the optimum flight parameters and flight patterns are determined by taking into consideration the GPS satellite constellation quality and availability, terrain, vegetation conditions, and accuracy requirements. Locations for differential GPS ground stations and placement of ground control points are also calculated based on flight geometry.

Since accurate LIDAR data acquisition technology depends heavily on GPS, it is imperative to obtain the most accurate GPS solution possible. To ensure the maximum accuracy, it is necessary to have a network of GPS differential ground stations throughout the project site so that no LIDAR data is collected more than 25 km from an operating GPS base station.

Flight Lines

Straight, parallel flight lines are strongly preferred with adjacent flight lines flown in opposite directions. Large turns with the airborne sensor cannot be used for data collection; however, data can be merged from multiple straight flight lines with different bearings. The coordinates of tracked flight line paths must be provided with the LIDAR data, including flight altitude, flight speed, and angle of scanning.

Coverage

Coverage must be complete. Areas not within two times the posting of data points are data voids. Except within bodies of water, raw data voids cannot exceed 5 percent of the collected area. In addition, the data must be collected without gaps wider than 2 pixels and empty areas more that 25 pixels. The areas of water bodies are an exception because LIDAR return may not occur if laser pulses are absorbed by the water.

Raw Data Collection

Point Density

The point density (shots/m^2 or shots/pixel) for the collected LIDAR data must be determined from the minimum statistics for single coverage (Figure 65), not from the average statistics that include overlaps.

Adherence to Flight Plan

Adherence to pre-planned flight lines must be carefully controlled during data acquisition to avoid incomplete coverage and data voids (Figure 66).

Figure 65. Distribution of LIDAR shots. Vertical strips show data swaths according to direction of flight lines. Water bodies and dark roofs have no reflections. Light colored strips show overlapped LIDAR data. The absence of LIDAR returns from dark roofs shows low sensitivity of LIDAR equipment.

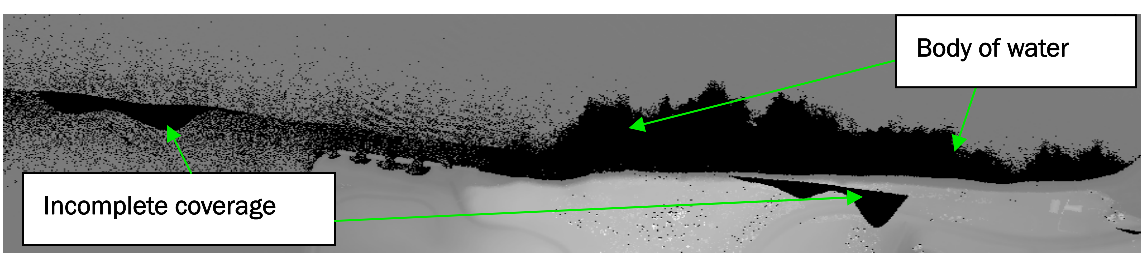


Figure 66. LIDAR data with two types of data voids due to bodies of water and incomplete coverage between two flight lines.

Mission Timing

Optimal mission times are determined using GPS positional dilution of precision (PDOP) and geometry charts, local weather and predicted solar activity. Collection must be accomplished during the most favorable season for LIDAR data collection; for example, leaf-off conditions are essential if a bare-earth digital terrain model is desired. Water surface elevations should be at normal levels for rivers and streams and tide-influenced bodies of water should be at low tide unless the mission is critical

Format of LIDAR data

The preferable format for the delivery of LIDAR data is:

- ASCII format – one line consists of information about one LIDAR shot (for both returns)
- Each file consists of one flight line – delivery of flight sessions is possible if all calibrations are good
- File of metadata for collected data for the project (coordinate systems, time of collection, etc.)
- File of metadata for each delivered flight line (coordinates of a flight line, etc.)
- The file name includes the geographic coordinates in the left bottom corner of tile.

Post-Processing LIDAR Data

Airborne GPS

The first step in post-processing LIDAR data is to analyze the airborne GPS files to produce a highly accurate GPS track of the aircraft. Any GPS data outside the tolerance window for the required accuracy is re-flown. If the data meets the accuracy specifications, the GPS track information is imported into the data processing software and used to process the x,y,z data points. Two base stations or more should be used for post processing. Both forward and reverse trajectory processing, using integers fixed at different times with different base satellites, is performed and compared to ensure good repeatability. Additional comparisons are made between different base stations to verify the accuracy of the GPS trajectories. Analysis of the chosen optimal trajectory is fully documented. This will include tables or charts that indicate the number, list, and elevation of satellites used. Documentation will also include PDOP, Vertical Dilution of Position (VDOP), standard deviation, and cycle slips.

Initial Accuracy Check

The processed data is passed on to quality review specialists who conduct spatial analysis on the laser data. The results are checked against the geodetic control points and ground truth checkpoints in open terrain to confirm that they are within the fundamental vertical accuracy specifications (see Section 0). The specialists also identify any data anomalies, gaps or other features that may affect data quality or accuracy, and submit those areas to be re-flown. At this point, the data is ready for vegetation/building removal and the creation of additional mapping products. The first and last returns should be nearly the same in open terrain. First and last returns are never merged to determine bare earth elevations.

Elevation Steps

Elevation steps or “jumps” in heights between flight lines must be corrected to not exceed the criteria specified in Section 0. Sharp elevation steps between adjoining strips must be avoided. If the data needs to be adjusted to fit adjoining flight lines, full documentation is required explaining why the data is to be shifted and by how much. Software is preferred to compare adjoining flight lines during processing.

Intensity Return Contrast

Calibration of reflection intensities must be performed to achieve small contrast (<10 percent) between flight lines (Figure 67) excluding water areas.

If smoothing heights or calibration of intensity between flight lines cannot be performed effectively by the original data provider, the data must be delivered for additional processing (by a LIDAR post-processing specialty firm) as a set of separated flight lines.

Bare-Earth Processing

The next step is to apply automated and/or manual procedures for removal of buildings and vegetation to generate a bare-earth Digital Terrain Model (DTM). Artifacts must be removed to reflect the true bare-earth terrain. Automated processing must not be overly aggressive in removing legitimate topographic features such as steep stream banks.

Figure 67. Merging two flight lines with poor calibration of intensity between flight lines creates three levels of intensity on the imagery-high, low, and mixed intensities.

Hydro-Enforcement

Some LIDAR datasets require hydro-enforcement to ensure the downward flow of water. This requires three basic steps:

1. Bridges and culverts are "cut" beneath the elevation of the water surface so that water will flow through a hydraulic model of the terrain rather than be blocked by a bridge or culvert that appears to be a dam.
2. Shorelines of lakes and reservoirs are leveled.
3. Stream centerlines (for small streams) and dual shorelines (for larger streams) are modeled with 3-D break lines that decrease uniformly from upstream elevations to downstream elevations so as to circumvent the undulations occurring naturally in LIDAR datasets that capture boulders, piles of dirt, rotting logs, and other features with variable elevations.

Summary of Recommended Data Requirements for Future Projects

Tables 8 and 9 list recommended data requirements for future projects.

Table 8. Quantitative requirements for LIDAR data delivery.

The LIDAR shall measure first and last returns for >99.9% pulses.
Number of Z-artifacts: <1/10,000,000 for negative Z-artifacts (<-30 cm) in usual geographic areas   <1/1,000,000 for negative Z-artifacts (<-30 cm) in for urban and similar areas   <1/1,000,000 for positive Z-artifacts (>30 cm)

Z-accuracy for 1.5 m LIDAR point spacing (2-m resolution of grid data) at the 95% confidence level:
    absolute RMSE = 15 cm
    relative RMSE = 5–7 cm (between LIDAR points in the same flight line)
    relative RMSE = 7–10 cm (between LIDAR points in the overlapping flight lines)
Z-accuracy for 0.75 m LIDAR point spacing (1-m resolution of grid data) at the 95% confidence level:
    absolute RMSE = 10 cm
    relative RMSE = 3–5 cm (between LIDAR points in the same flight line)
    relative RMSE = 5–7 cm (between LIDAR points in the overlapping flight lines)

X,Y accuracy: 1-m or better at the 95% confidence level
LIDAR data is not collected more than 25 km from an operating GPS base station.
The data should be collected without gaps wider than 2 pixels and empty areas more that 25 pixel (except for areas of water bodies).
The LIDAR Point Density shall be such that there is at least one measurement per 1-m grid cell (pixel) for 95% of the area except for bodies of water. Data voids or “holiday” areas, defined in Section 0, less than 25 pixels are permitted if those holiday areas do not exceed 2 pixels in width for all non-hydrographic areas.
Contrast of intensity returns between flight lines must be smaller than 10%.

Table 9. Additional requirements for LIDAR data collection.

The LIDAR system shall provide intensity measurements.
The LIDAR system shall measure returns from dark roofs, asphalt, and electric wires.
The leaf-off vegetation condition is strongly preferable for generation of DTM and building models.
Straight, parallel flight lines are strongly preferred. Large turns with the airborne sensor cannot be used for data collection.
Collection of LIDAR data in urban areas should compensate for “shadowing” in urban canyons.

Collected data should be delivered in:
    ASCII format, one line = one LIDAR shot (with both returns)
    File of metadata for collected data

Factors That Affect Costs and DTM Quality

In the LIDAR industry, the cost for the collection and processing of LIDAR data is proportional to the data collection area (per km^2, per sq mi or per tile with specific size) and highly dependent on customer requirements and data product deliverables.

The most important factors that affect LIDAR data and DTM cost are:

- area of collection (both size and shape)
- density of data
- required accuracy of bare-earth data
- location of project area (proximity to airport and usable survey control points)
- level of processing to minimize artifacts (automated vs. manual processing)
- availability of ground truth checkpoints
- requirement for hydro-enforcement.

Figure 68 shows a rough estimation of the cost of LIDAR and DTM data collection in relation to the size of area of collection and density data or required accuracy DTM.

The cost of collecting 30-cm data for small area can reach several thousand dollars per km^2. The cost of standard 2-m resolution LIDAR data varies from \$100.00 to \$300.00 per km^2 for a relatively large area, about 1,000 km^2.

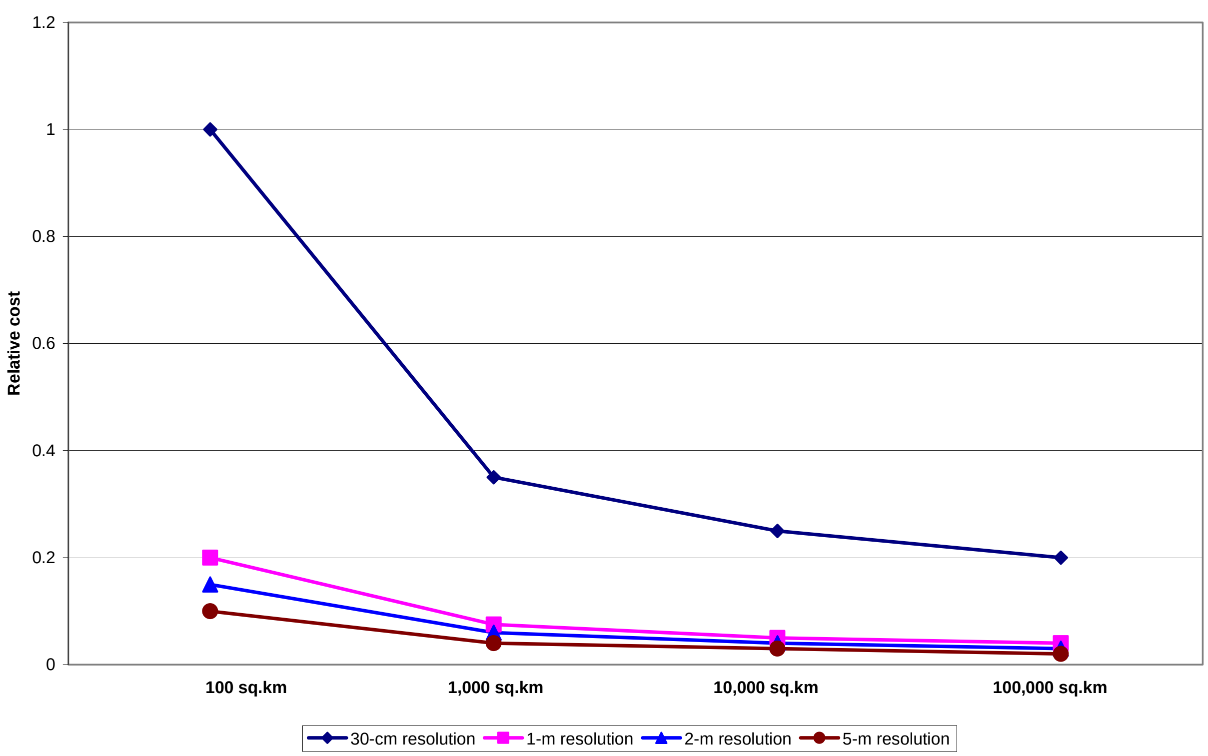


Figure 68. Estimation of typical cost of LIDAR data and DTM based on the size of the collection area and required DTM resolution.

The most important factors that affect LIDAR data and DTM accuracy are:

- short baselines for airborne GPS control
- low flying heights above mean terrain
- narrow scan angles to use data closer to nadir
- density of LIDAR data points
- density and canopy closure of vegetation
- number of artifacts in LIDAR data
- accurate ground control points
- leaf-off conditions
- shallow slopes
- minimum water areas.

The data in Table 10 assume a double coverage of LIDAR data or 50 percent overlapping flight lines. Fifty percent overlapping means that light colored areas with larger statistics will cover the entire area; the area of single coverage will be negligibly small.

Table 10. Relationship between the densities of LIDAR points, resolution of data, and accuracy of models.

	30-cm resolution	1-m resolution	2-m resolution	5-m resolution
Spacing of LIDAR data	0.23 m	0.75 m	1.5 m	3–4 m
Accuracy of DTM	0.3 m	1 m	2 m	5 m
Intensity imagery	0.15–0.3 m	0.5–1 m	1–2 m	2.5–5 m
Polyline for roads	0.3–0.6 m	1–2 m	2–4 m	5–10 m*
Polyline for water	0.6 m	2 m	4 m	10 m*
Polyline for buildings	0.6 m	2 m	4 m	10 m*
*Data with 5-m resolution has restricted applicability to the extraction of road, small bodies of water, and buildings. Many objects, like residential buildings smaller than 10 m will be lost.				

Figure 69 shows a rough estimate of the relative costs for generating different models for four typical geographic areas: DTM, vegetation, buildings, bodies of water, and roads.

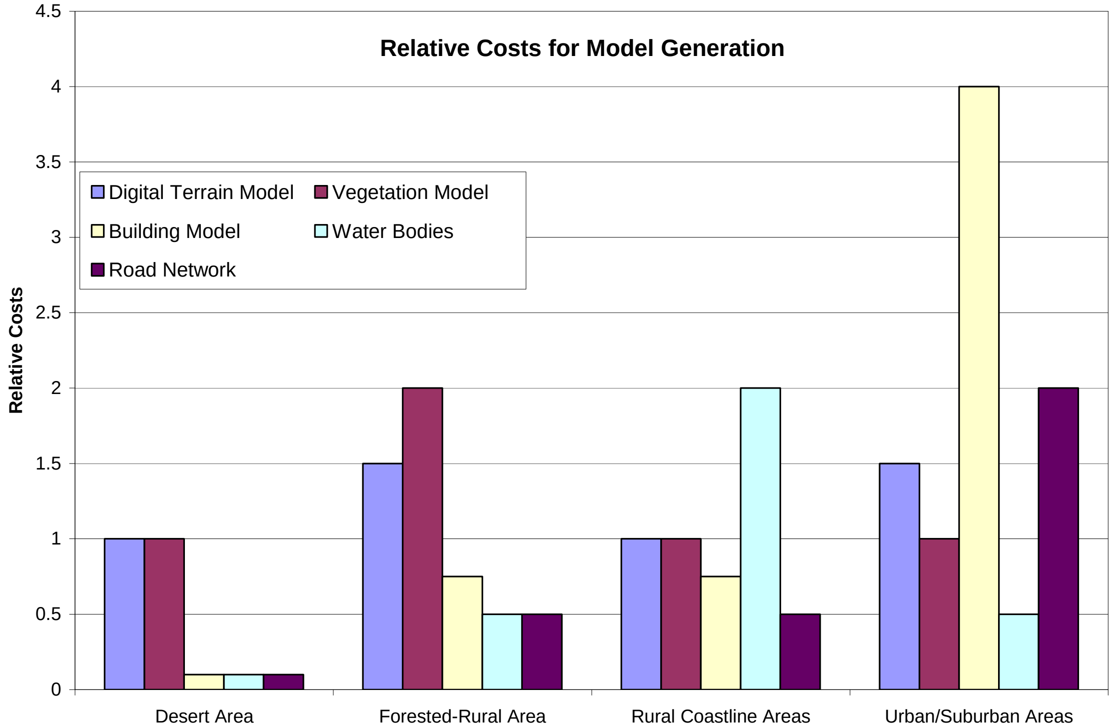


Figure 69. Relative costs for generation of models for four typical geographic areas: bare earth, vegetation, buildings, bodies of water, and roads.

The industry standard for the generation of high quality DTM is one third of the price of the collection of LIDAR data. Most companies that collect LIDAR data use COTS processing and DTM generation software that cannot address many artifacts in data and cannot meet different customer requirements. The best results, especially for special customer requests or for problematic areas, are achieved by contracting with a company that employ a team of specialists and improves its proprietary software for each project.

The cost of intensity imagery is approximately 12–15 percent of the cost of the DTM. The total price for the generation of a complete set of models is comparable to the cost of collecting LIDAR data, or for urban or other complex areas, several times more.

Collection and Quality

LIDAR data must be collected:

- According to the LIDAR collection requirements document.
- By a proven, experienced LIDAR data vendor
- With a set of necessary control points and field tests

- With permanent, independent expert control available immediately after collection and delivery to ensure data quality. Delivery cannot be accepted without expert positive analysis.
- Under an agreement that includes re-processing, re-delivery or/and re-collection of LIDAR data if the experts feel the information is unsatisfactory.

Price

The cost of LIDAR data varied from vendor to vendor. Some vendors offer an unrealistically low price for the collection and processing of LIDAR data; however, the quality of the LIDAR data they provide is out spec in many cases. The customer may discover the situation many months after the start of a project. This is not good for the customer, the contracting company and the LIDAR industry. An attractive price must be weighed against the reputation of the firm providing the bid.

LIDAR Data from Ground Vehicles

LIDAR sensors can be mounted on ground vehicles to collect 3-D surfaces around a vehicle, rather than looking downward from an aircraft. The requirements for collecting and processing ground vehicle LIDAR data are similar to the requirements for airborne LIDAR systems. The most important differences are:

- better resolution and quality due to smaller distances
- less expensive per number of shots
- increased time to cover the project area due to the lower vehicle speed
- collection area is restricted due to shadowing and access restriction.

Processing LIDAR data from ground sensor presents more challenges due to the complex 3-D geometry of collected data. Appendix A to this report includes samples of two ground systems that illustrate the capabilities of a typical ground-based LIDAR system.

LIDAR Data Quality Assessment, Independent Validation, and Verification

Although each LIDAR data provider performs in-house Quality Assessment and Quality Control (QA/QC), the results are normally biased and suspect. Independent Validation and Verification (IV&V) is best performed by a firm that specializes in LIDAR standards and independent QA/QC of LIDAR data. Such IV&V is both quantitative and qualitative.

Quantitative QA/QC

In 2004, the American Society for Photogrammetry and Remote Sensing (ASPRS) endorsed the LIDAR accuracy assessment standards originally proposed by the National Digital Elevation Program (NDEP). These standards use a minimum of 20 surveyed checkpoints in each major land cover category representative of the area, and include three specific methods:

1. Fundamental Vertical Accuracy (FVA) computes vertical accuracy at the 95 percent confidence level in open terrain only (dirt, sand, short grass) based on RMSEz x 1.9600. The RMSE method is applicable only in open terrain where elevation errors follow a normal distribution. For example, if elevations are expected to be comparable to 2 ft contours, the RMSEz should equal 18.5 cm, and the FVA should equal 36.3 cm (14.3 inches) at the 95 percent confidence level, based on RMSEz x 1.9600. The FVA determines the performance of the LIDAR sensor, prior to the introduction of post-processing algorithms for vegetation removal.
2. Supplemental Vertical Accuracy (SVA) computes vertical accuracy at the 95 percent confidence level separately for each major land cover category based on the 95th percentile error for each land cover category, e.g., bare earth, weeds and crops, scrub, forests, and built-up areas. The SVA determines the performance of the LIDAR system including post-processing algorithms. In vegetated areas, LIDAR bare earth elevation errors do not necessarily follow a normal error distribution because potential systematic errors might cause bare earth elevation errors to all be positive; if vegetation and buildings are not removed, for example.
3. Consolidated Vertical Accuracy (CVA) computer vertical accuracy at the 95 percent confidence level for all land cover categories combined, based on the 95th percentile error for all checkpoints combined. The RMSE method is not appropriate because errors do not necessarily follow a normal error distribution. The CVA provides a single statistic for the overall accuracy of the LIDAR dataset in all land cover categories.

Qualitative QA/QC

Qualitative QA/QC is best performed by a firm that specializes in independent QA/QC of LIDAR data. This includes manual evaluation of hill shades to determine artifacts, seams between flight lines, and systematic errors that lead to poor contour lines when DTMs are converted into contours. This typically results from the lack of hydro-enforcement needed for contours to correctly “behave” in the vicinity of streams; however, it also results from the lack of break lines needed to define distinctive features. Some firms “oversmooth” the data to minimize artifacts and thereby

eliminate terrain undulations and special features that should be retained in the DTM. Surveyed cross sections help to identify stream channel geometry and evaluate whether or not the LIDAR dataset has been "over-smoothed."

Imagery Data Collection Requirement and Quality Verification

Data from Multi-Spectral and Hyperspectral Remote Sensors of Aerial Vehicles

Commercial mapping firms operate a variety of airborne multispectral and hyperspectral sensors capable of automated feature identification and data extraction. Generally, multispectral sensors with 3–4 wide bands can only perform Anderson Level 1 classifications (Table 11) of major land cover categories using ERDAS software for example. Hyperspectral sensors with perhaps a hundred narrow bands can perform Level 2 or 3 classifications with far greater specificity, using ENVI software for example. However, hyperspectral sensors are ideal for automated feature identification primarily when images are classified with ground truth field samples from the same image scenes. Variations in sun angles, soil moisture content, and other variables generally limit the usefulness of hyperspectral data libraries that store "hyperspectral signatures" acquired under conditions that might not be duplicated when imagery is reacquired on a different date.

Data from Multi-Spectral and Hyperspectral Remote Sensors of Ground Vehicles

Multi-spectral and hyperspectral data from sensors of ground vehicles have two advantages over airborne. Ground vehicle collected data has better resolution and complex 3-D geometry, while airborne data has 2-D geometry. The ability for spectral classification of extracted objects is similar for both ground vehicle and airborne data.

Table 11. Land use and land cover classification system for use with remote sensor data.

Level I	Level II
Urban or Built-up Land	1.1 Residential 1.2 Commercial and Services 1.3 Industrial 1.4 Transportation, Communications, and Utilities 1.5 Industrial and Commercial Complexes 1.6 Mixed Urban or Built-up Land 1.7 Other Urban or Built-up Land
Agricultural Land	2.1 Cropland and Pasture 2.2 Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas 2.3 Confined Feeding Operations 2.4 Other Agricultural Land
Rangeland	3.1 Herbaceous Rangeland 3.2 Shrub and Brush Rangeland 3.3 Mixed Rangeland
Forest Land	4.1 Deciduous Forest Land 4.2 Evergreen Forest Land 4.3 Mixed Forest Land
Water	5.1 Streams and Canals 5.2 Lakes 5.3 Reservoirs 5.4 Bays and Estuaries
Wetland	6.1 Forested Wetland 6.2 Nonforested Wetland
Barren Land	7.1 Dry Salt Flats. 7.2 Beaches 7.3 Sandy Areas other than Beaches 7.4 Bare Exposed Rock 7.5 Strip Mines Quarries, and Gravel Pits 7.6 Transitional Areas 7.7 Mixed Barren Land
Tundra	8.1 Shrub and Brush Tundra 8.2 Herbaceous Tundra 8.3 Bare Ground Tundra 8.4 Wet Tundra 8.5 Mixed Tundra
Perennial Snow or Ice	9.1 Perennial Snowfields 9.2 Glaciers

References

1. Optech Inc., http://www.optech.ca/

2. Martin Flood (ed.). 24 May 2004. *ASPRS Guidelines. Vertical Accuracy reporting for LIDAR data*. ASPRS Lidar Committee.

3. FEMA. Appendix 4B. Airborne Light Detection and Ranging Systems. *Flood Hazard Mapping*, http://www.fema.gov/fhm/lidar_4b.shtm

4. James R. Anderson, Ernest E. Hardy, John T. Roach, and Richard E. Witmer. 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper 964. A revision of the land use classification system as presented in U.S. Geological Survey Circular 671.

5. Multi-Resolution Land Characteristics Consortium. Webpage, http://www.mrlc.gov/index.asp

6 Generation of Synthetic Scenes and Digital Terrain Models Using LIDAR Data from White Sands Missile Range

Description of Data Source

Map and Geographical Type of Project Area.

Figures 70 and 71 show the project area, a 10 x10 km parcel located in White Sands, Otero County, NM as DTE4 COA1.

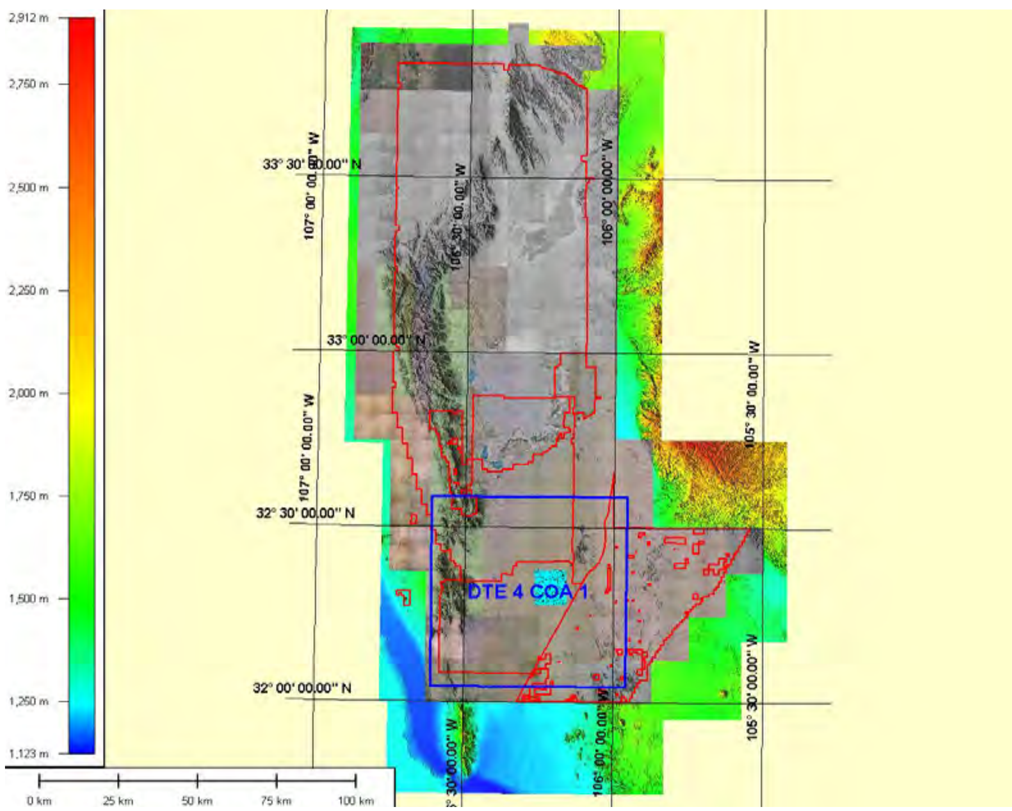


Figure 70. The project area (DTE4 COA1) is outlined in blue.

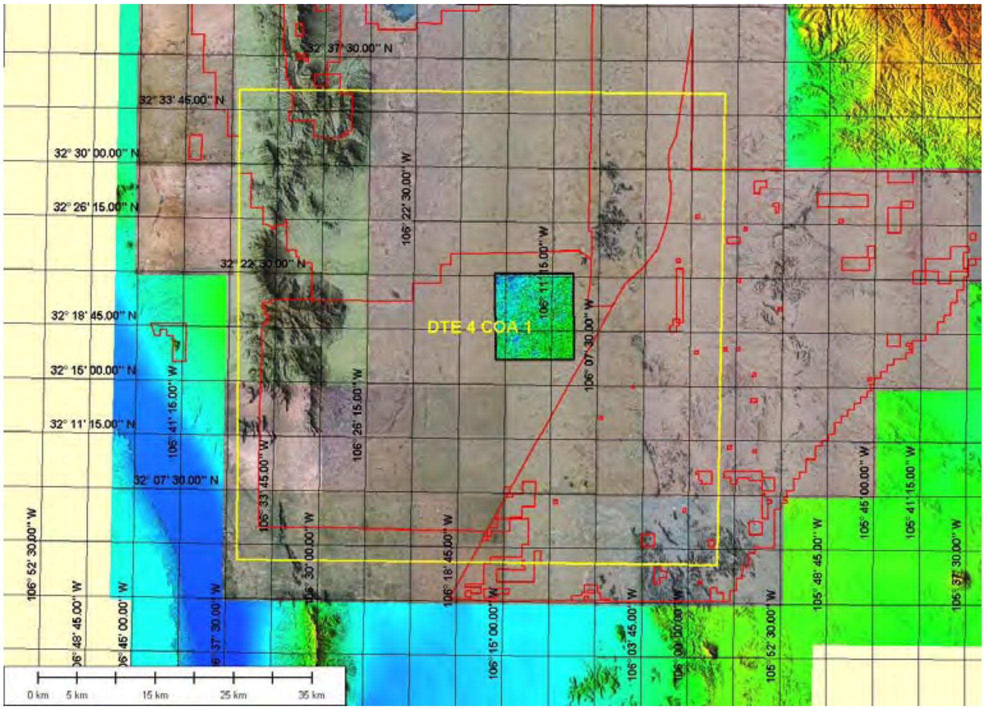


Figure 71. Detailed map with project area (DTE4 COA1).

The project area is a sand desert with small hills, covered by bushes and tall grass. Figures 72 through 75 show the various landscape characteristics found in the project area.


Figure 72. More than 50 percent of the project area is covered by dense bushes and tall grass.

Figure 73. LIDAR data was collected when the vegetation was in leaf-on time. Average height of the vegetation is 2–3 ft.

Figure 74. Sand hills approximate 2 m in height.

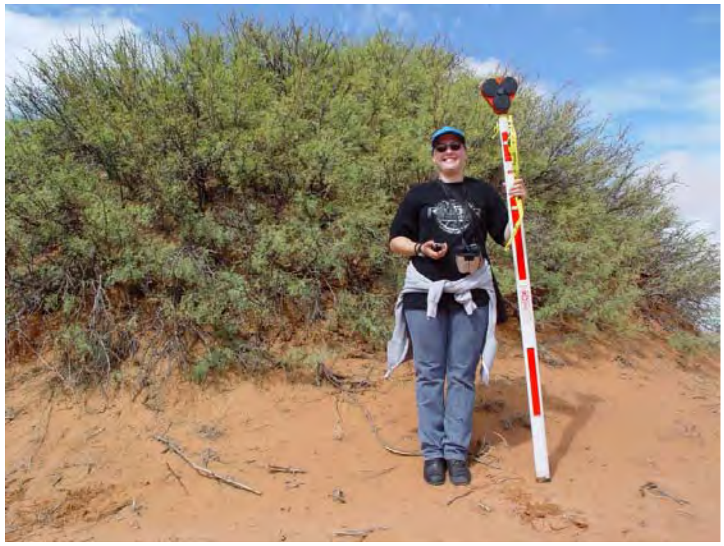


Figure 75. A typical hill with bushes.

Collected LIDAR data

From an airborne platform on White Sands Missile Range, LIDAR data was collected for a 100-km² area with 1-m spatial resolution, and for a 10–12 km² area with 30-cm resolution. (Figures 76 through 78 show coverage maps.) Data was delivered to the GDATP/GIST team via 1024 x 1024 m tiles that overlapped by 24 m on the right and bottom edges.

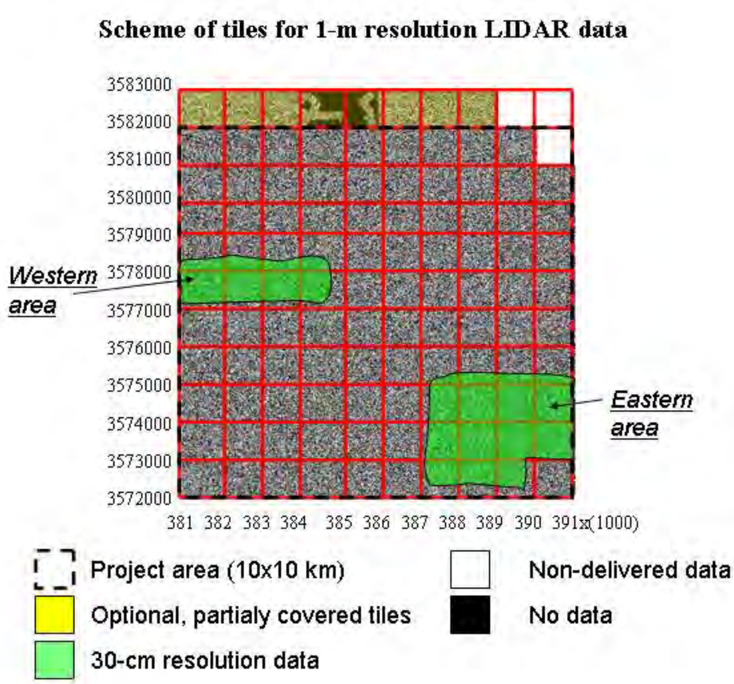


Figure 76. Ninety-nine tiles inside the project area were covered completely by 1-m resolution data, eight tiles outside project area covered partially; 107 processed tiles were delivered for 1-m resolution data.

Detailed distribution of 30-cm resolution LIDAR data by tiles in the western and eastern areas is shown in Figures 77 and 78; 30-cm LIDAR data was processed using sub-tiles sized 341 x 341-m. Large 1,024 x 1,024 m delivered tiles were divided into nine sub-tiles without overlapping.

An important characteristic of the collected LIDAR data is the single last return. Only 0.1 percent of the data included first returns for vegetation higher than 6 ft. Single return data can seriously underestimate the volume of vegetation. The question of possible underestimation of vegetation must be investigated during field research and compared to the representation of real vegetation in collected LIDAR data.

Accuracy and Artifacts of LIDAR data

Artifacts are regions of anomalous elevations, oscillations, or ripples within the Digital Elevation Model (DEM)/vegetation data resulting from malfunctioning sensors, poorly calibrated instrumentation, and other sources of systematic errors during collection of LIDAR data.

During the processing of LIDAR data collected by the Government at White Sands Missile Range, the team identified several serious artifacts in collected LIDAR data that affected the overall quality of the delivered data products and models:

- negative (below ground) strip-like artifacts
- negative point-like artifacts
- positive (above ground) strip-like artifacts
- positive point-like artifacts
- calibration artifacts with different intensities of reflection in different flight lines.

Usually such artifacts concentrate near specific flight lines and reflect errors in the collection of LIDAR data. Figures 79 through 86 show different artifacts. (Each shows minimal heights for all LIDAR shots in each pixel.)

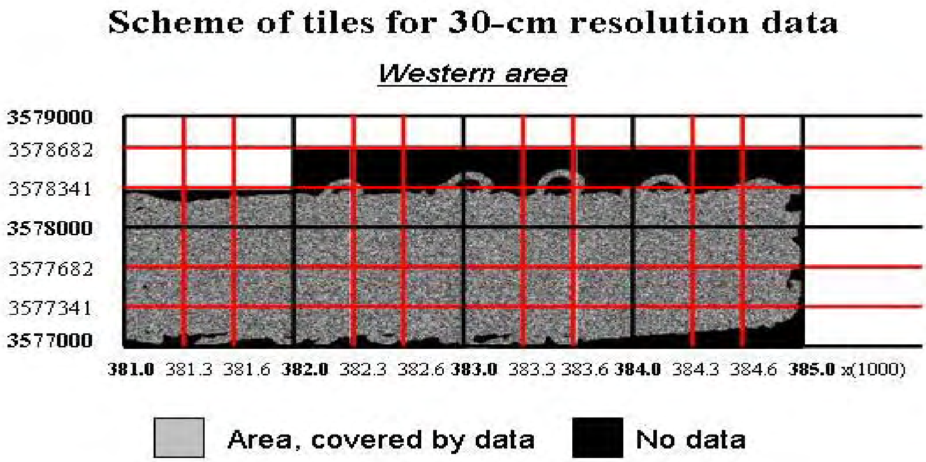


Figure 77. 22 sub-tiles were covered completely and 35 sub-tiles covered partially. The quality of LIDAR data in the central project area is better than on the edges due to greater statistics and a fewer number of artifacts; 51 processed sub-tiles were delivered for this area; six almost empty sub-tiles were ignored.

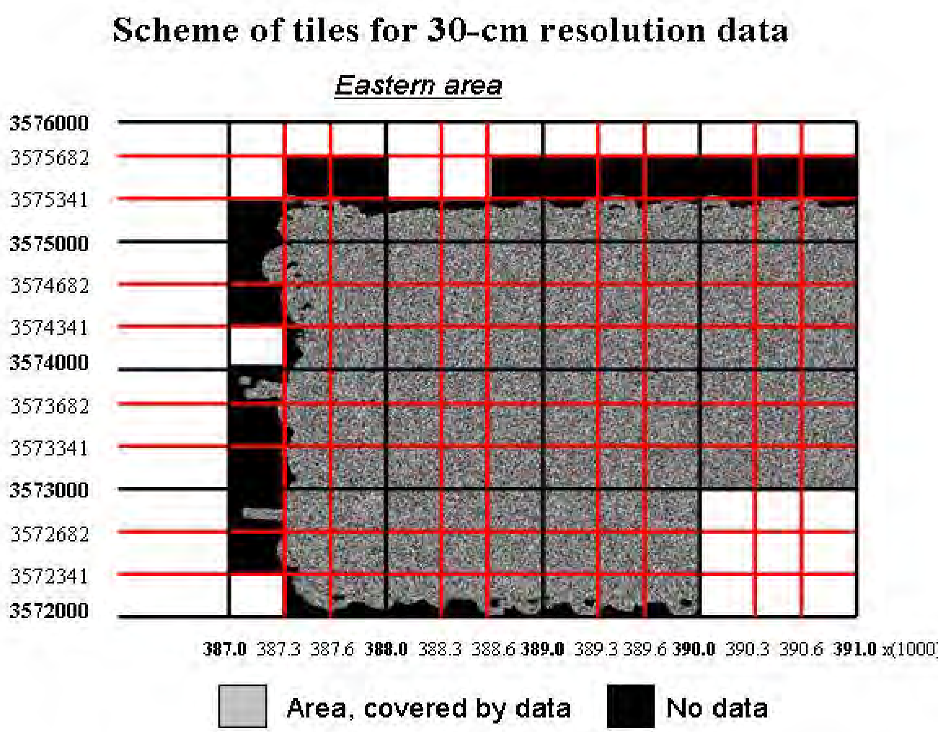


Figure 78. The eastern area of collected LIDAR data with 30-cm resolution; 74 sub-tiles were covered completely, 44 sub-tiles partially; 104 processed sub-tiles were delivered for this area; 14 almost empty sub-tiles were ignored.

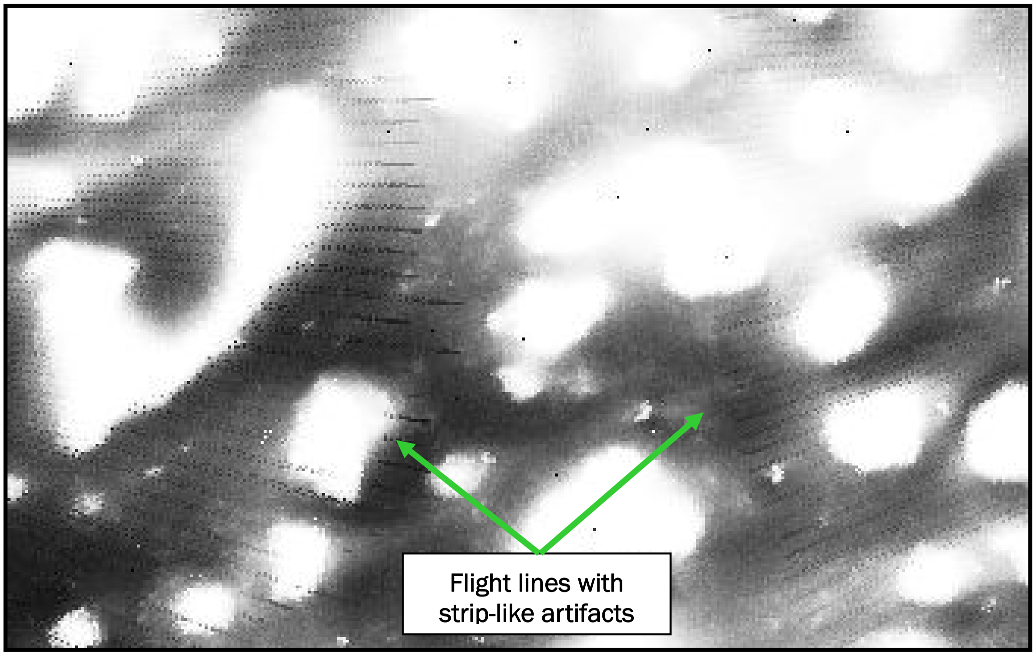


Figure 79. Flight lines with strip-like artifacts, underground and above-ground “corn rows,” from 30-cm LIDAR data found in sub-tile 389000_3574341.

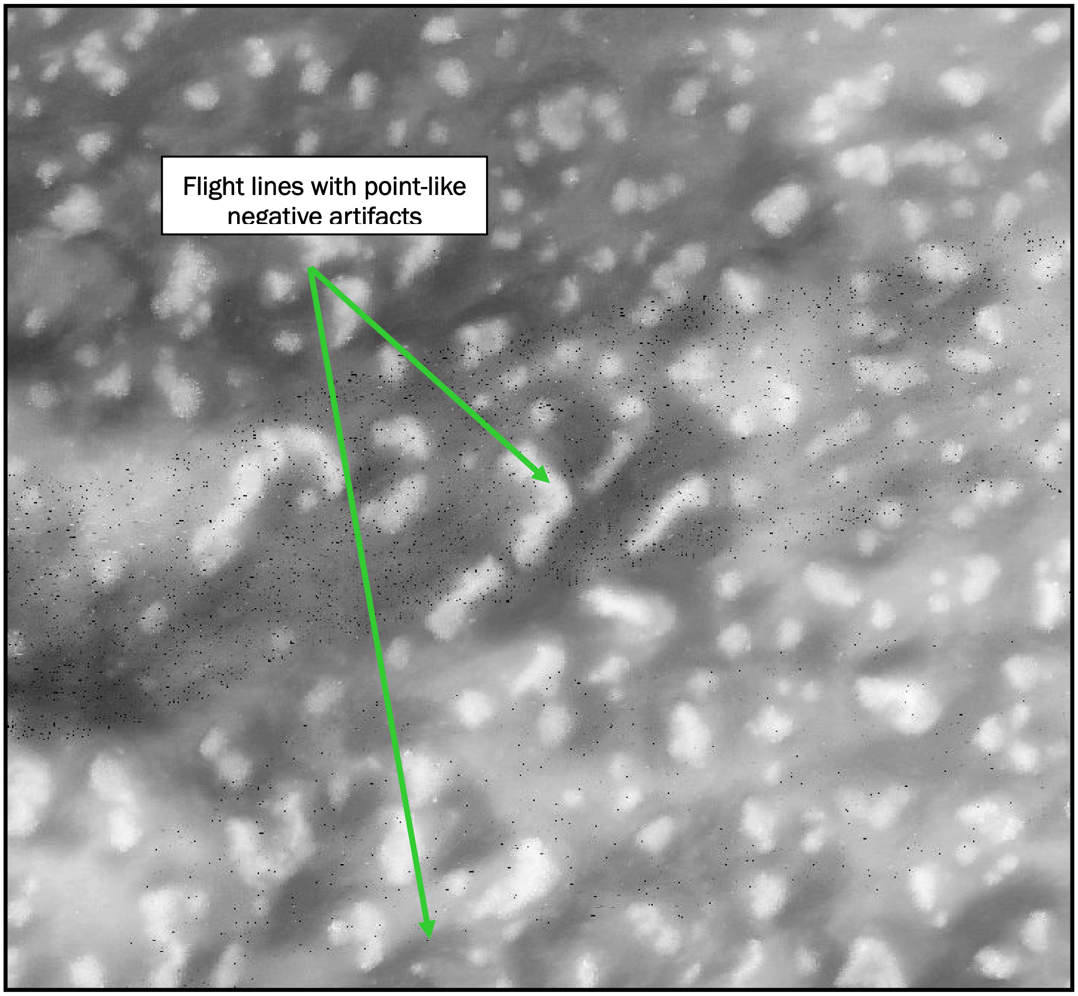


Figure 80. Flight lines with excessive point-like negative artifacts (underground reflections) in 30-cm LIDAR data from sub-tile 381341_3577341.

In high quality commercially available LIDAR data, artificial underground reflections must be smaller than 1 artifact per 10 million LIDAR shots. In the presented LIDAR data, the number of artifacts is at least 3–4 times larger. In a typical geographic area, similar underground artifacts cannot be filtered effectively; consequently, the quality of the Digital Terrain Model (DTM) will be very low. Figures 82 and 83 show the effects of negative artifacts to the DTM. For this specific desert area with small vegetation, more aggressive approaches can be used for filtering artifacts (described in Section 0). These aggressive approaches can be applied to improving DTM generation; however, they are ineffective for improving models of vegetation that were heavily damaged by the positive artifacts in collected LIDAR data (Figures 83 and 84).

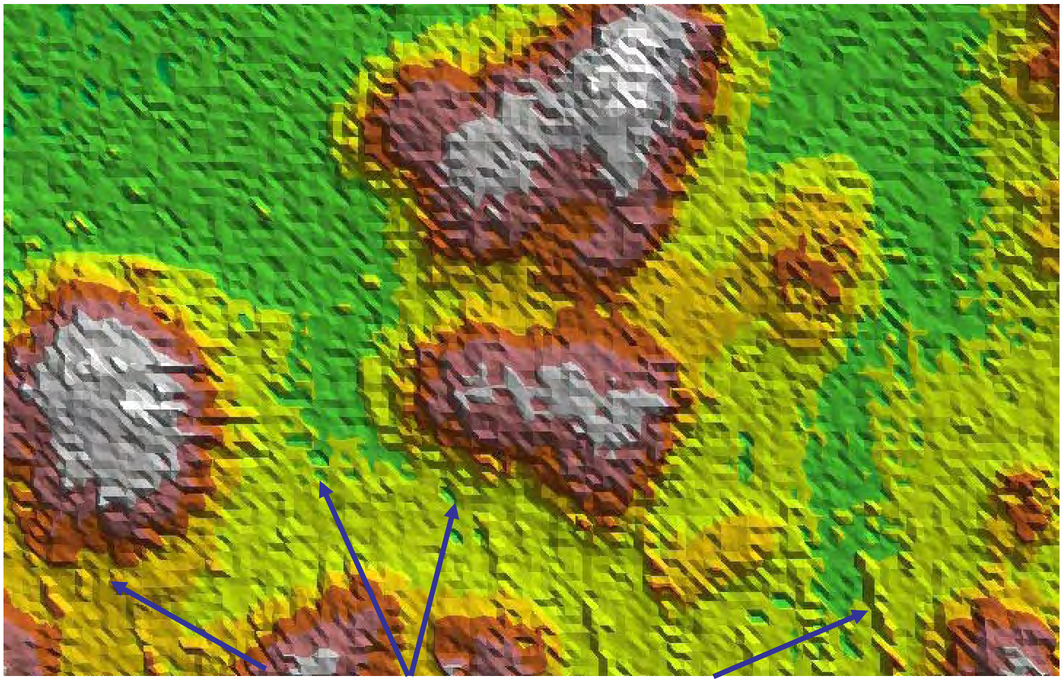


Figure 81. Three dimensional view of strip-like and point-like artifacts in Z_min imagery.

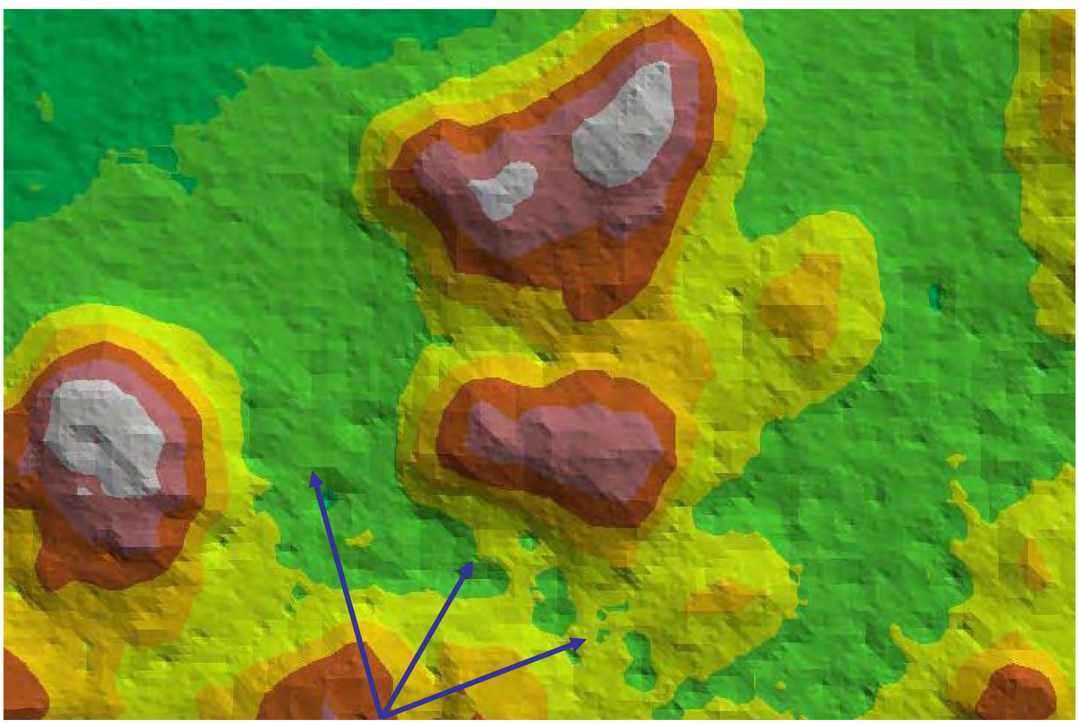


Figure 82. Three dimensional view of the effect of numerous negative point-like artifacts to the DTM with positive strip-like artifacts filtered out.

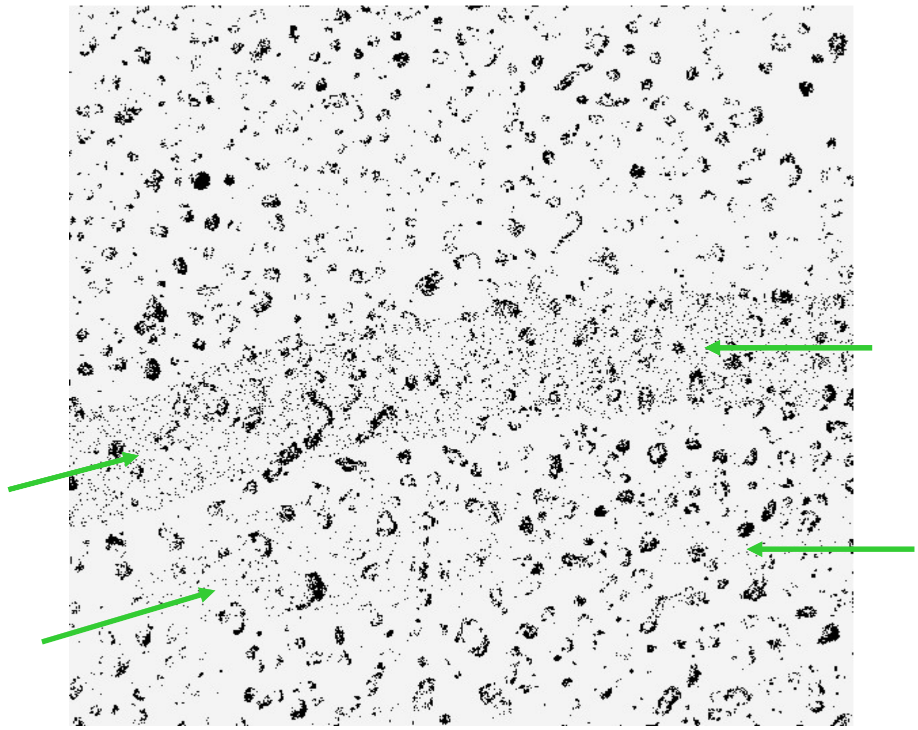


Figure 83. Abnormal flight lines in 30-cm LIDAR data in sub-tile 381341_3577341 have excessive point-like artifacts, both negative and positive. Bands of positive artifacts show relative vegetation heights >15 cm.

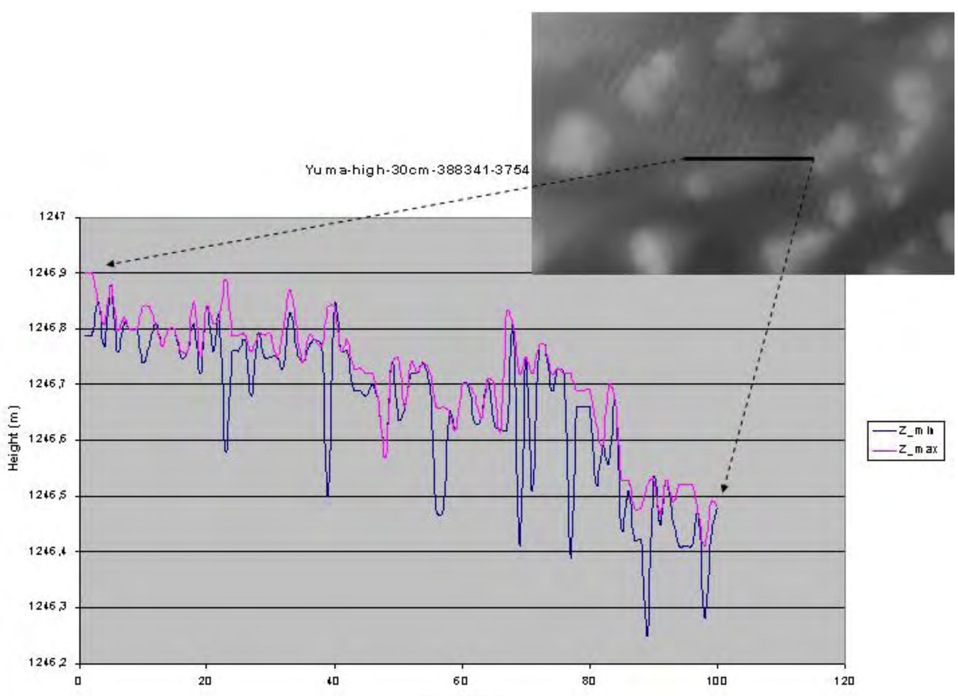


Figure 84. Cross section of a strip-like artifact in sub-tile 388341_3754000 reveals that this artifact has both positive and negative components with amplitude up to 1 foot.

Figure 85 shows one example of poor calibration between intensity of reflection in different flight lines.

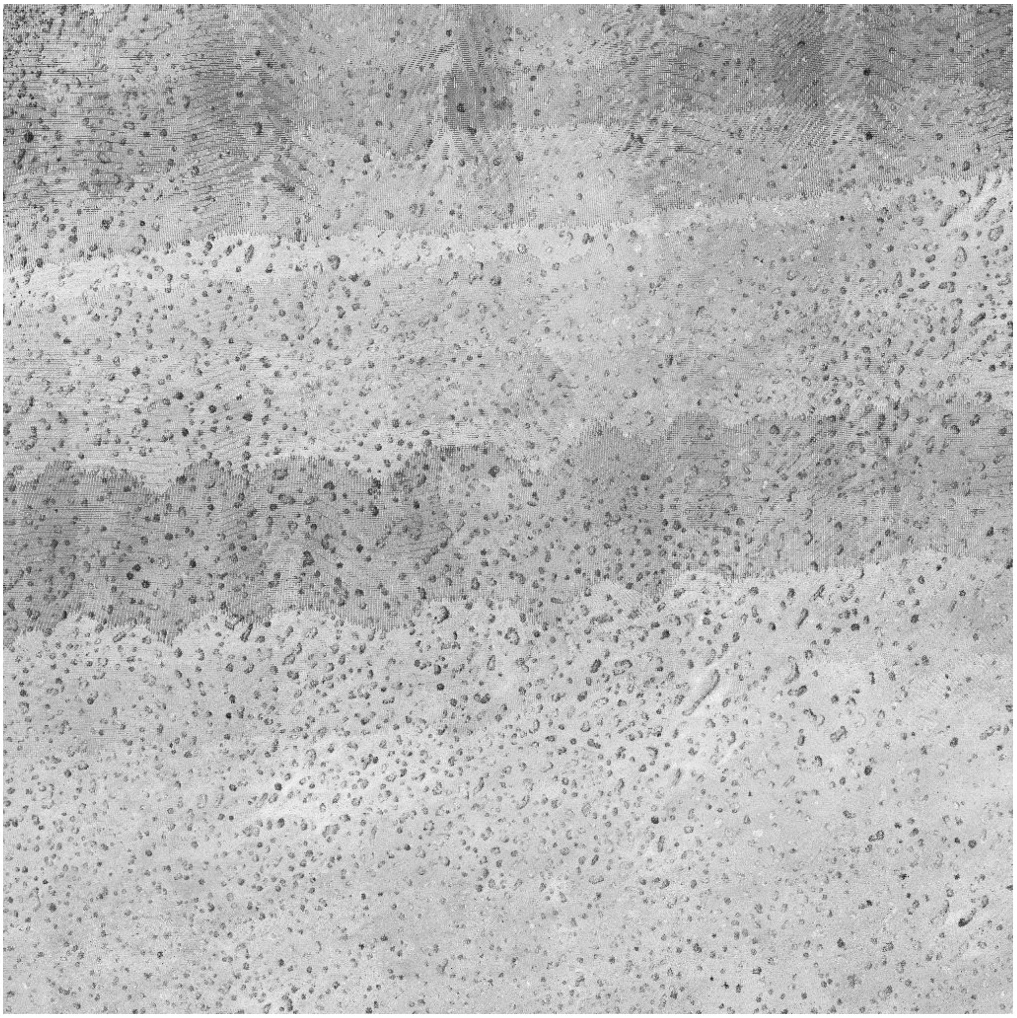


Figure 85. An example of poor calibration of intensity returns between different flight lines that affects data about vegetation reflectivity (1-m resolution data for tile 381_3577).

The accuracy of collected LIDAR data cannot be compared with control points because they were not provided. Consequently, it was only possible to use non-direct methods for estimation of accuracy. Figure 86 shows strong correlation between two independent sets of LIDAR data (30-cm and 1-m resolutions) inside 10–20 cm, and it is possible to estimate that the accuracy of LIDAR data will be close to these numbers.

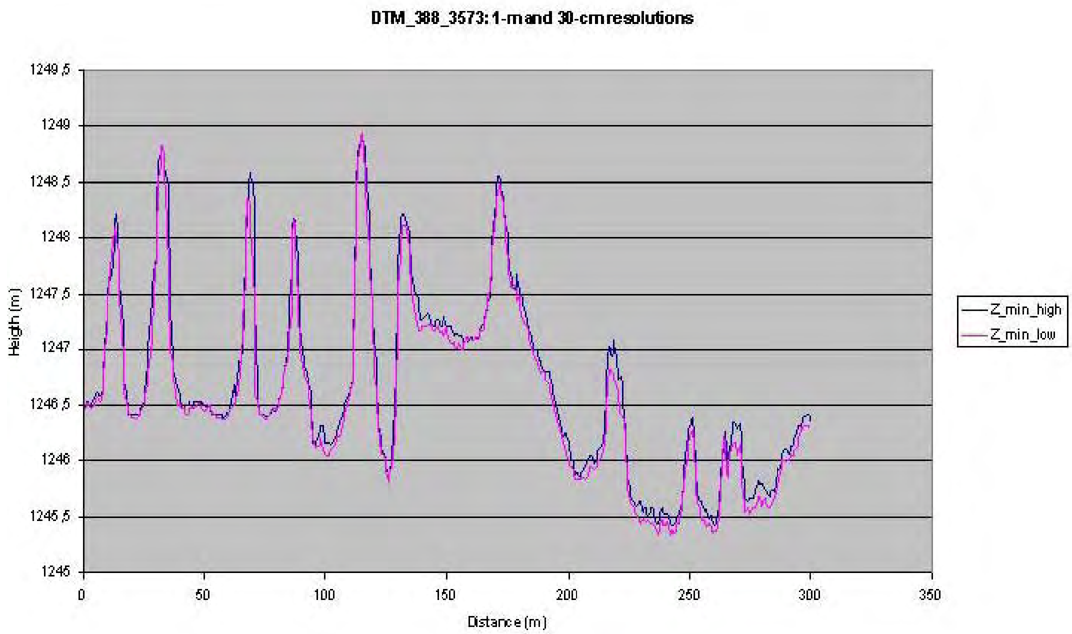


Figure 86. Profiles for Z_minimal for two independent sets of collected LIDAR data: 1-m (Z_min_low) and 30-cm (Z_min_high) resolutions demonstrate strong correlation.

Methods of Processing LIDAR data

Artifacts Filtering

To eliminate the positive and negative point-like and strip-like artifacts from DTM, the following approaches were used:

1. **Averaging height from last returns.** Figure 87 shows the effectiveness of this approach. An additional benefit of this method is that DTM from average heights is free from systematic underestimation of the height of bare earth, which is typical for DTM generated from minimal heights. This approach cannot be applied to forested regions due to serious error in DTM.

Z_minimal and Z_average for part of tile of 1m_389_3576

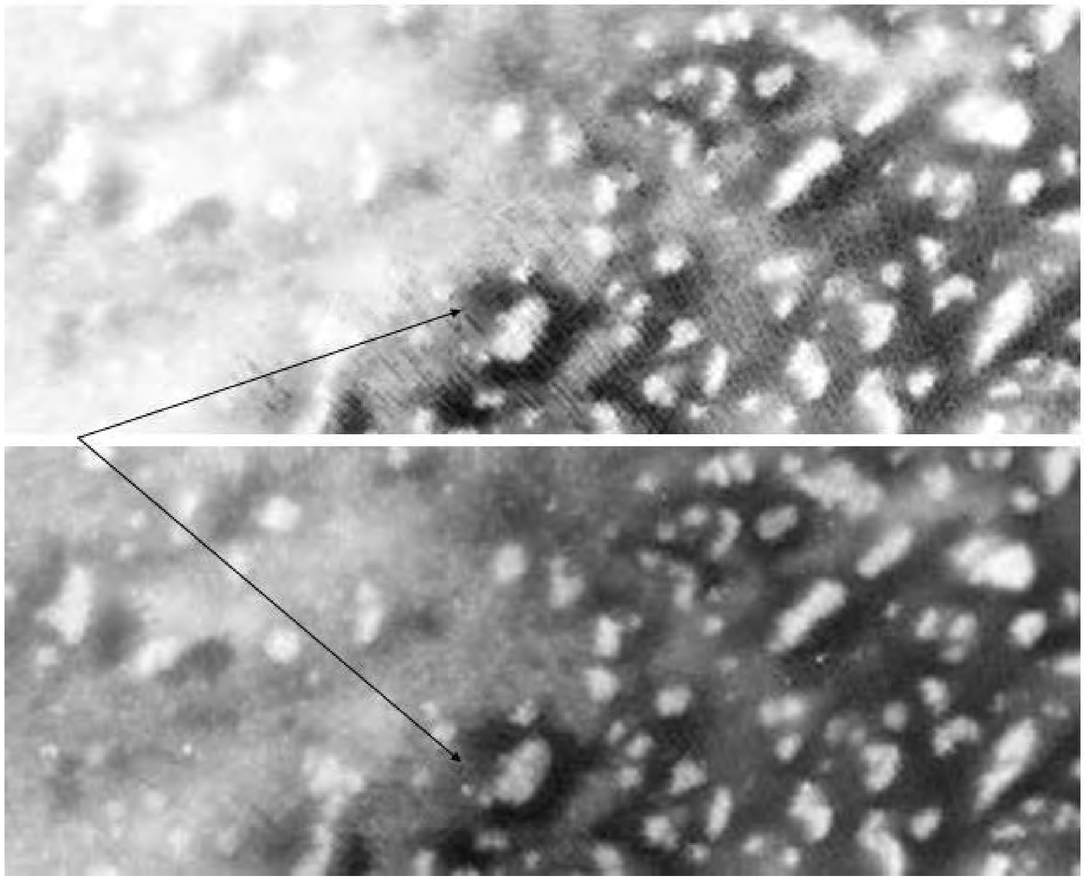
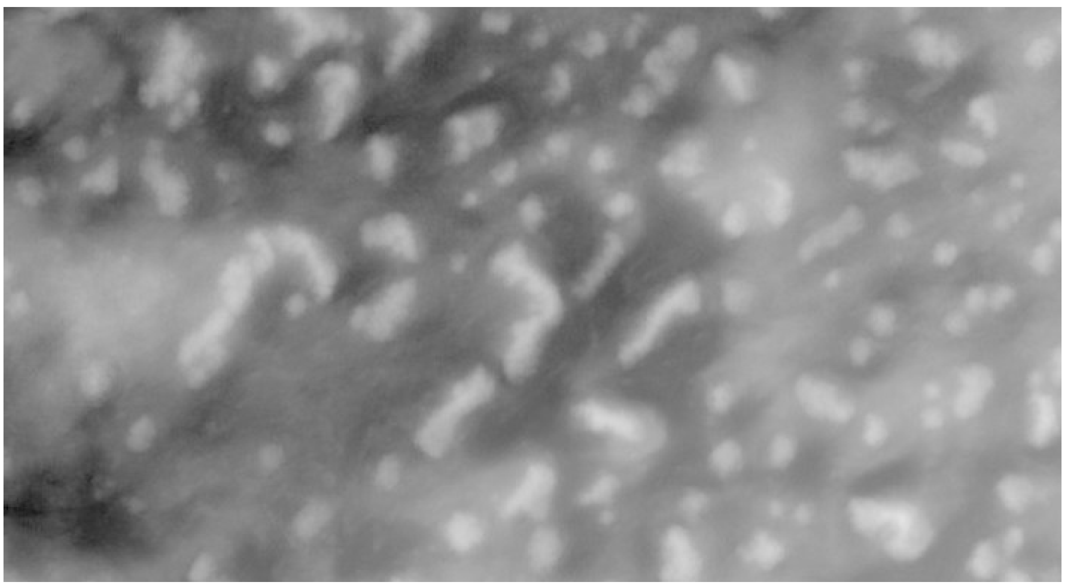


Figure 87. The top photograph shows strip-like artifacts in minimal height surface. In the bottom photograph, these artifacts are eliminated by averaging of heights.

2. **Averaging height to the nearest neighbor.** This procedure also eliminates approximately 90 percent of point-like artifacts. Other parts of point-like artifacts are separated out by an additional filter that compares heights of holes and holidays to neighboring pixels. Figure 88 shows the effectiveness of this procedure for filtering point-like artifacts in DTM; practically 100 percent of point-like artifacts were eliminated. This approach cannot be applied in a forested area because the deep penetrations of the laser beam, which are essential for DTM generation, are erased in forest.

(A)



(B)

Figure 88. A: An illustration of point-like artifacts in minimal height surface (tile 381341_3577341 of 30-cm resolution data); B: Artifacts are eliminated by averaging of heights and additional filtering.

These aggressive approaches cannot be used for improving models of vegetation from positive artifacts due to two concerns: (1) using of averaging heights for generation of vegetation models can eliminate some positive artifacts, such as large parts of the vegetation, that is not acceptable; and (2) there are no effective and secure filtering techniques for positive artifacts due to the fact that positive noise and vegetation have similar sizes and other characteristics. Any possible approaches to filtering the artifacts will delete part of vegetation. In most cases, that is not acceptable.

Filling Gaps

In this project, four types of interpolation of empty areas (see Appendix B "Description and Specification of Data Product") were used. Figure 89 shows the tile out project area with numerous empty areas. Figure 90 shows the result of different types of interpolation. For example, type of interpolation #3 is flat filling of gaps using the average height of the pixels on the boundary of the empty area because it was less than 2500 pixels. The type of interpolation #4 is was flat filled using the minimal value of the height of the pixels on the boundary of the empty area because it was larger than 2500 pixels. Areas interpolated by type 1 and 2, with high and middle accuracy, are invisible. Complete information regarding all interpolated pixels is located in the attributes or metric file (see Appendix B) for the estimation of errors in height of each pixel.

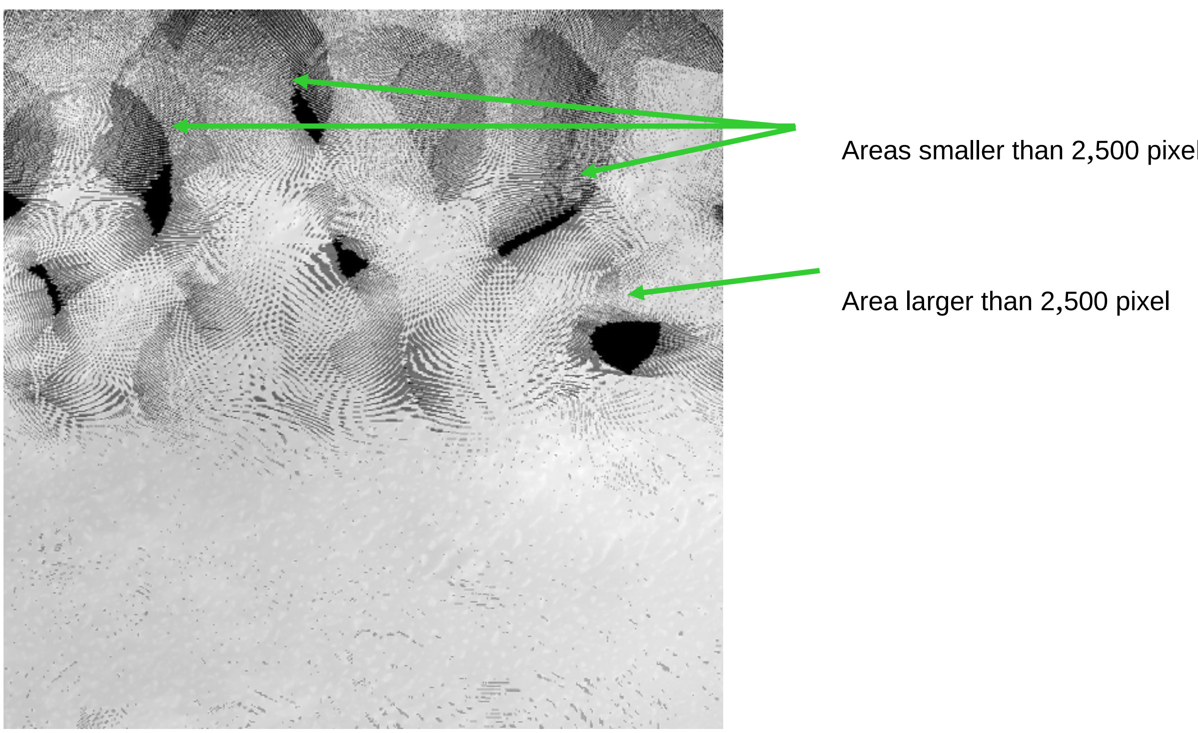


Figure 89. DTM for tile 381_3582 outside project area (1-m resolution) with empty areas due to lack of LIDAR data (dark points and areas).

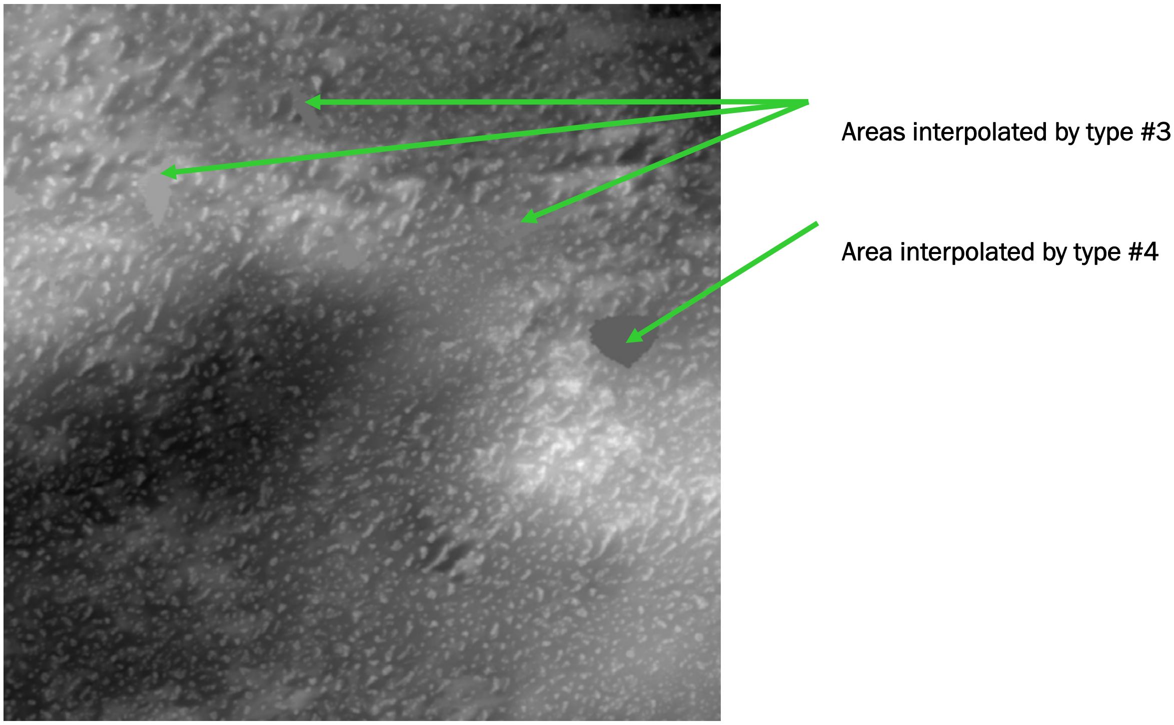


Figure 90. DTM after filling gaps by different types of interpolation.

Extraction of Foliage and DTM Generation

The Virtual Surfaces Method (VSM) for DTM generation and foliage filtering from LIDAR data was developed by the scientists at the Greenwich Institute for Science and Technology. The VSM is a local approach that uses analytical surfaces for local fitting earth surface. Using spherical or elliptical surfaces, independent of each other, VSM jointly creates a global surface. The process for foliage filtering is integrated with bare-earth surface interpolation in filtered pixels. The VSM possesses the benefits of popular methods for DTM generation, such as the Vosselman-Sithole and Kraus-Pfeifer techniques; however, VSM avoids most of the problems associated with these approaches (see details GDATP/GIST technical paper “Terrain Modeling: Mathematical Methods and Numerical Algorithms to Describe Terrain and Surface Features” 2004). Greater accuracy and other benefits of the VSM were tested during an extensive industrial application for processing LIDAR data and DTM generation in an approximately 20,000 km^2 area. After a comparison of generated DTM with original surface (or LIDAR data), it is possible to extract all objects above bare-earth: foliage, structure and noise (Figure 91). Noise covers almost all surfaces without foliage and presents instrumental noise (for example, LIDAR strips) and small details with a height smaller 15 cm.

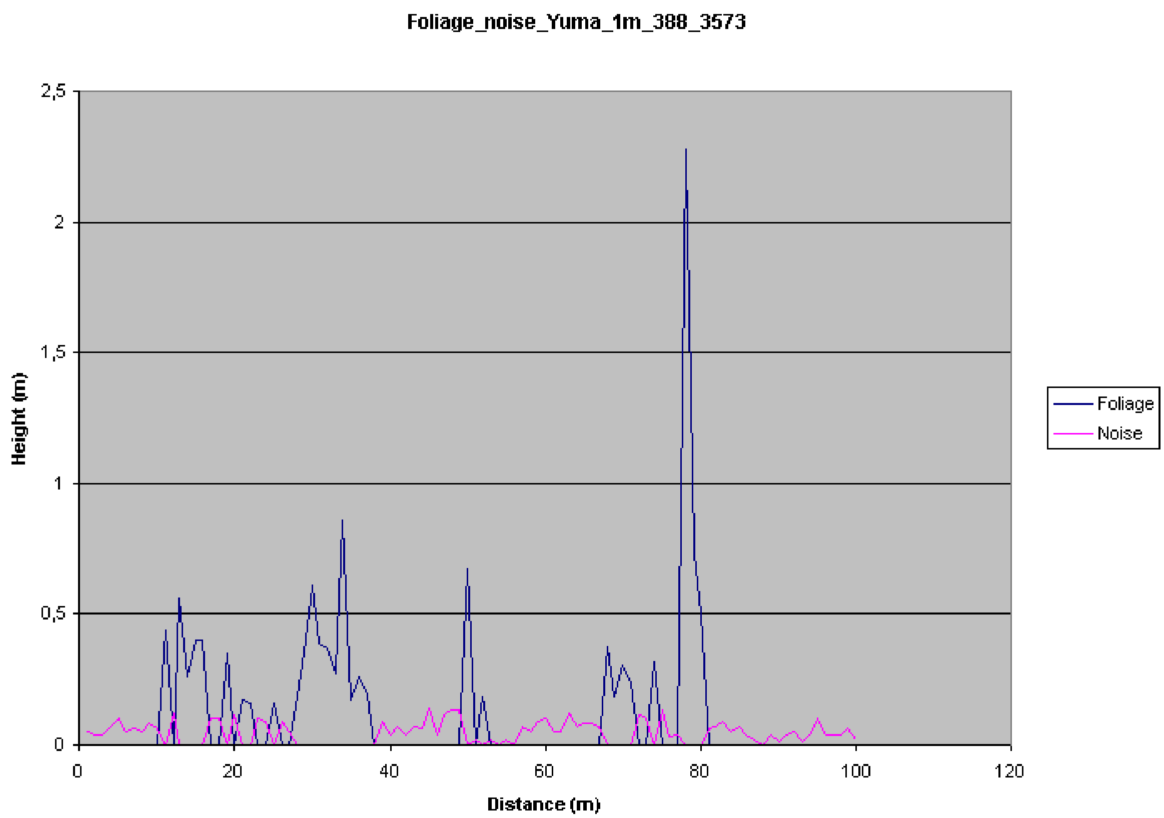


Figure 91. Cross-section for heights above DTM (for 1-m resolution data). LIDAR returns higher than 15 cm were classified as foliage; smaller heights were classified as noise.

Figures 92 and 93 show the results of processing by VSM. These figures also illustrate that noise has a strong correlation with clusters of vegetation and represents the lowest part of the vegetation.

Vegetation Modeling

To improve the presentation of foliage, a section of the highest noise (7.5–15 cm) pixels around preliminary classified vegetation pixels was considered as vegetation. Figure 94 shows the theory of extraction of vegetation as pixel with heights > 15 cm (black) and connected (by sides or corners) pixel with heights from 7.5 to 15 cm (gray). Clustering of vegetation is performed by separation of sets of pixels, which are connected by the sides only, not corners.

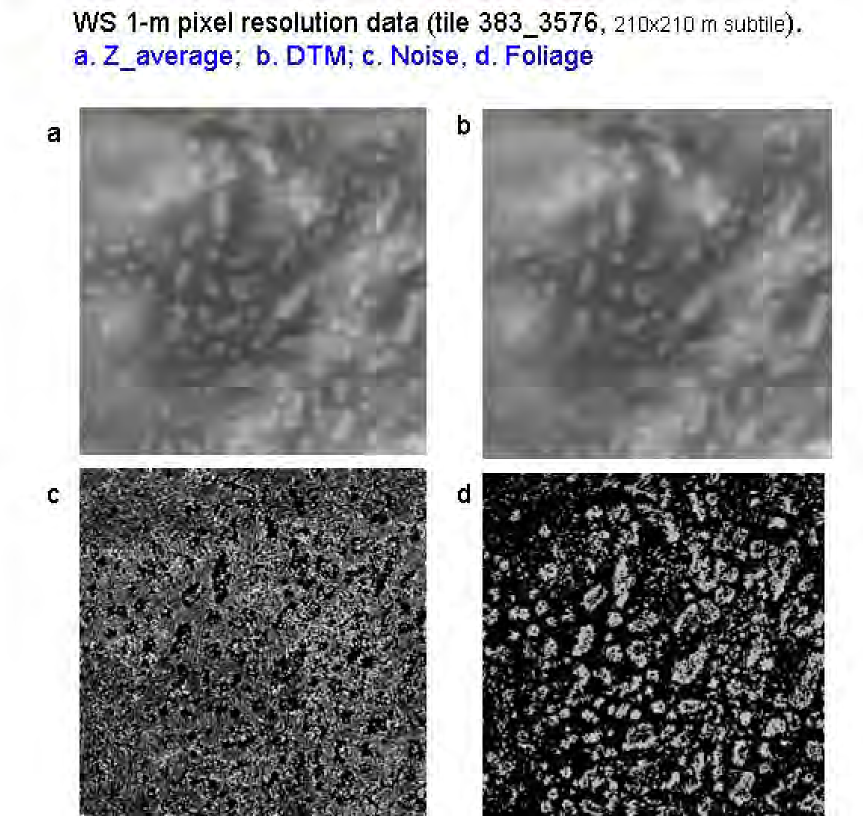


Figure 92. Maps of heights for original data (Z_average) and products of processing by VSM: DTM, noise (<15 cm) and foliage (>15 cm).

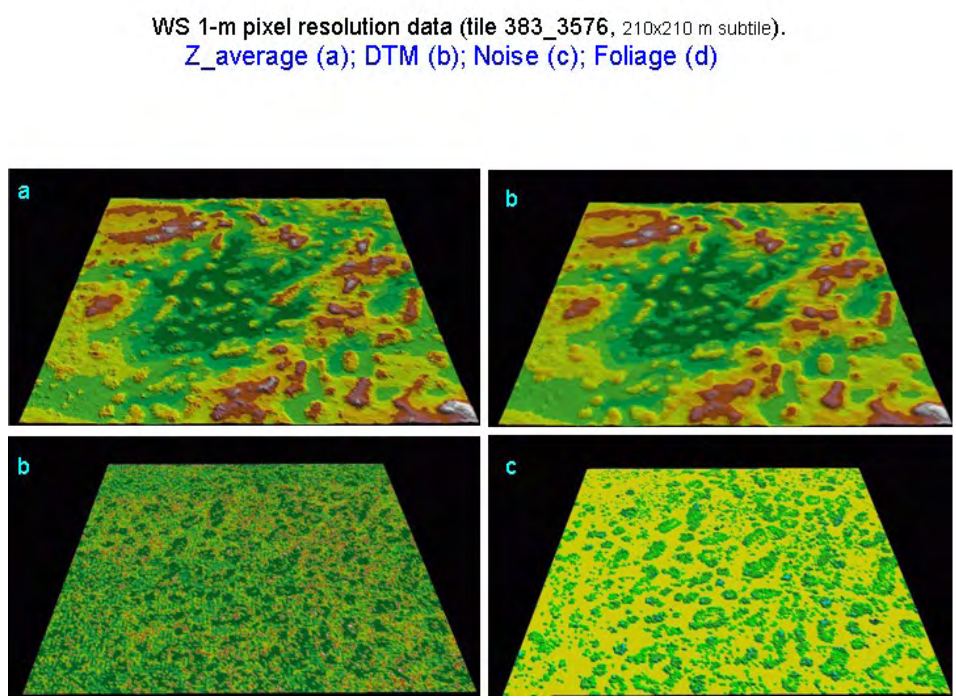


Figure 93. A three dimensional view of original data and products of processing from Fig. 92. The part with the highest noise (b) correlates with the vegetation on (c).

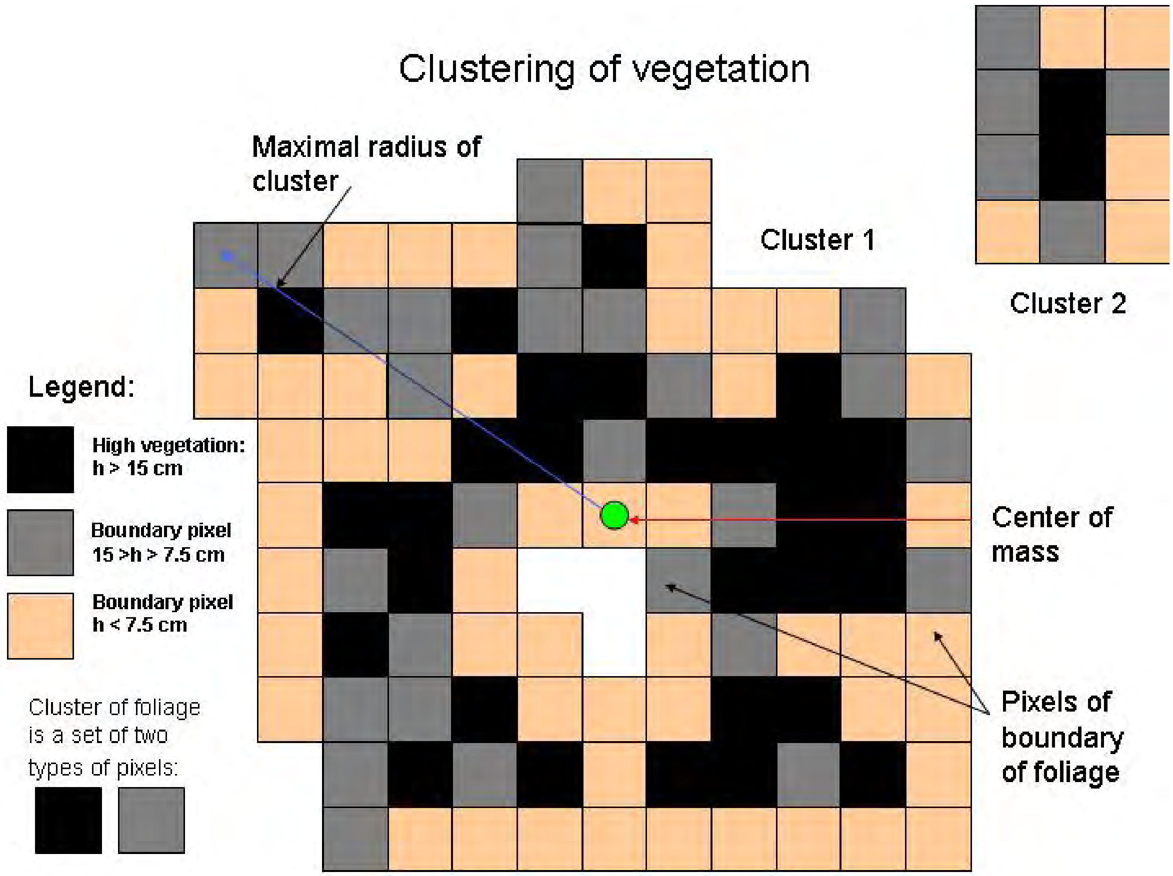


Figure 94. Ideology of classification of pixels of vegetation and clustering of vegetation.

Each cluster of foliage was delivered in two formats, local (or pixel level) and integral (or model level) (see Appendix B):

- *Pixel level*: Includes a list of each pixel and complete information about each pixel.
- *Model level*: Each cluster is represented by integral parameters and a geometric primitive (ellipsoid).

Each cluster of vegetation with absolute heights in each pixel was interpolated by the top hemisphere of 3D-ellipsoid with semi-axes $\mathbf{a,b,c_{max}}$, where $\mathbf{a,b}$ are the semi-axes of outside ellipse (see Figure 95); and $\mathbf{c_{max}}$ is the vertical axis of ellipsoid, which describes the maximal vertical extension of cluster. In addition, the vertical axis of ellipsoid that describes average vertical extension of cluster, $\mathbf{c_{aver}}$, was calculated. The absolute height of the center of an ellipsoid is the minimal height from the absolute heights of all pixels of vegetation in this cluster. This ellipsoid was considered the zero-order approximation for foliage cluster. In many cases, such model is too simple; it does not present the complex shape of the cluster correctly. An integral model (**vint** files in data delivery) provides a general

idea about the position of this cluster and its number in database. For this foliage cluster, one can extract a local model with the most accurate information on pixel-level presentation (**vloc** files in data delivery) from the database.

Figure 95 shows extracted clusters of vegetation (local model) and elements of integral models – ellipses contoured with each cluster of vegetation. In many cases, these ellipses overlapped and touched.

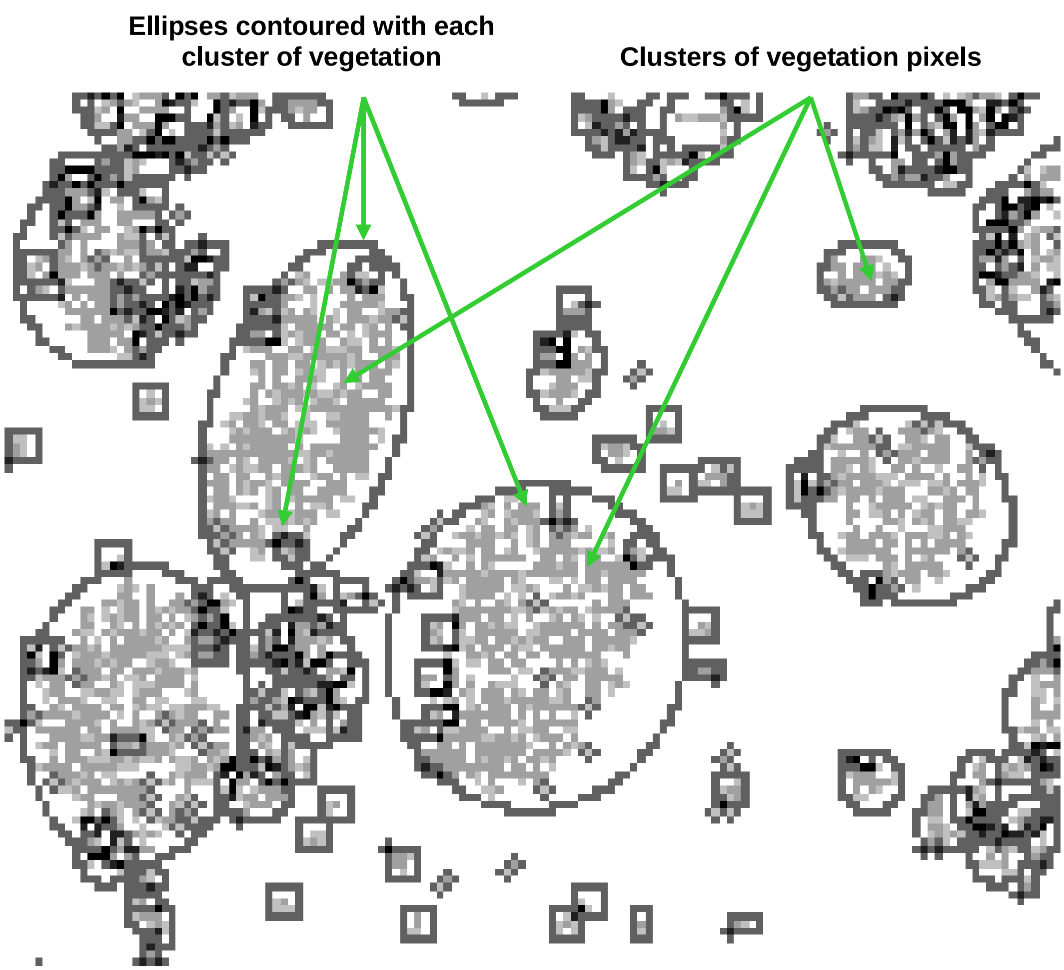


Figure 95. Extracted clusters of vegetation: dark pixels show high vegetation >15 cm, lighter pixel show boundary pixel with heights 7.5–15 cm. Ellipses (darkest pixel) represent integral models. Black pixels illustrate where an ellipse overlapped with a pixel from another cluster; an ellipse cannot overlap with pixels of its own cluster.

Samples of Data Product

DTM

Figure 96 shows the typical result of DTM generation at White Sands Missiles Range for 1-m resolution. Figure 97 shows DTM with 30-cm pixels.

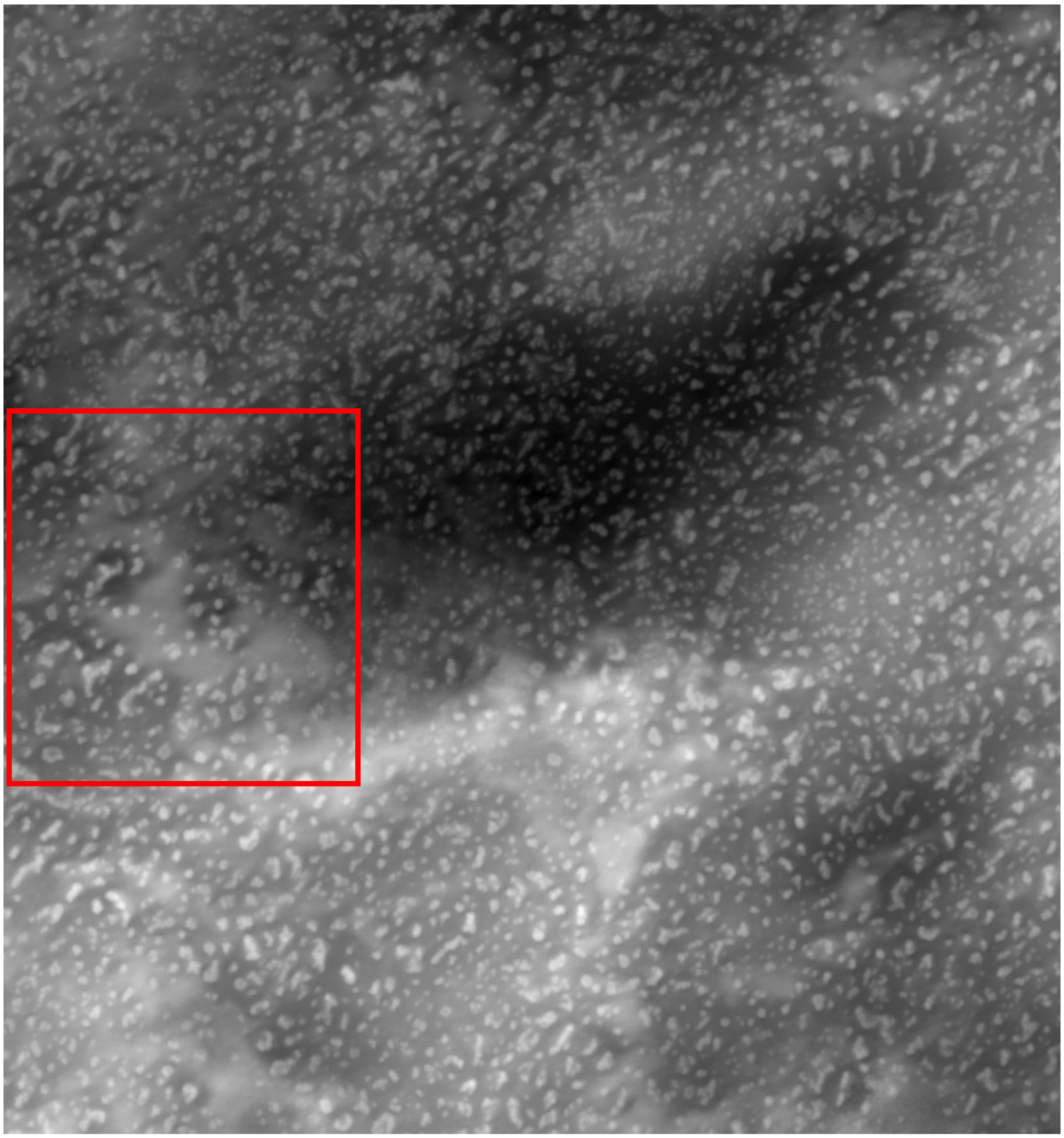


Figure 96. DTM (bare-earth) for tile 388000_3574000 (1-m resolution LIDAR data). Project area has two scales for relief: small-scale local hills connected with vegetation and large-scale variation of relief. Red box indicates the position of sub-tile 388000_3574341.

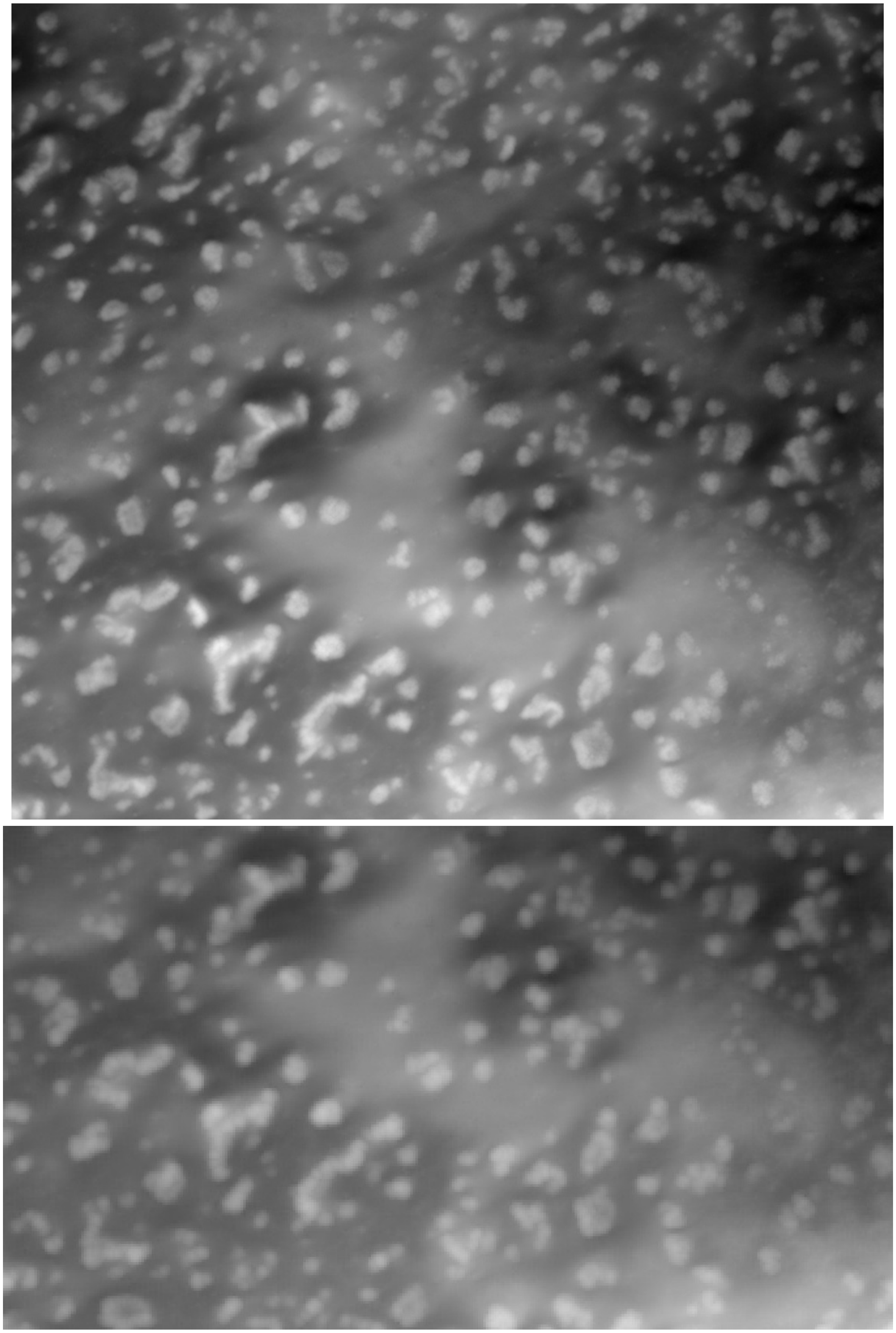


Figure 97. Top image presents the DTM (bare-earth) for sub-tile 388000_3574341 (30-cm resolution LIDAR data). Lower image shows a segment of DTM with 1-m resolution for bottom section of the sub-tile.

Intensity Imagery

Intensity imagery for typical LIDAR data returns approximate the information provided by aerial photography with a similar resolution (Figure 98). The collected LIDAR data calibration of the equipment, however, was not of good quality; many intensity images, especially in 30-cm resolution, are of a lower quality (Figure 99).

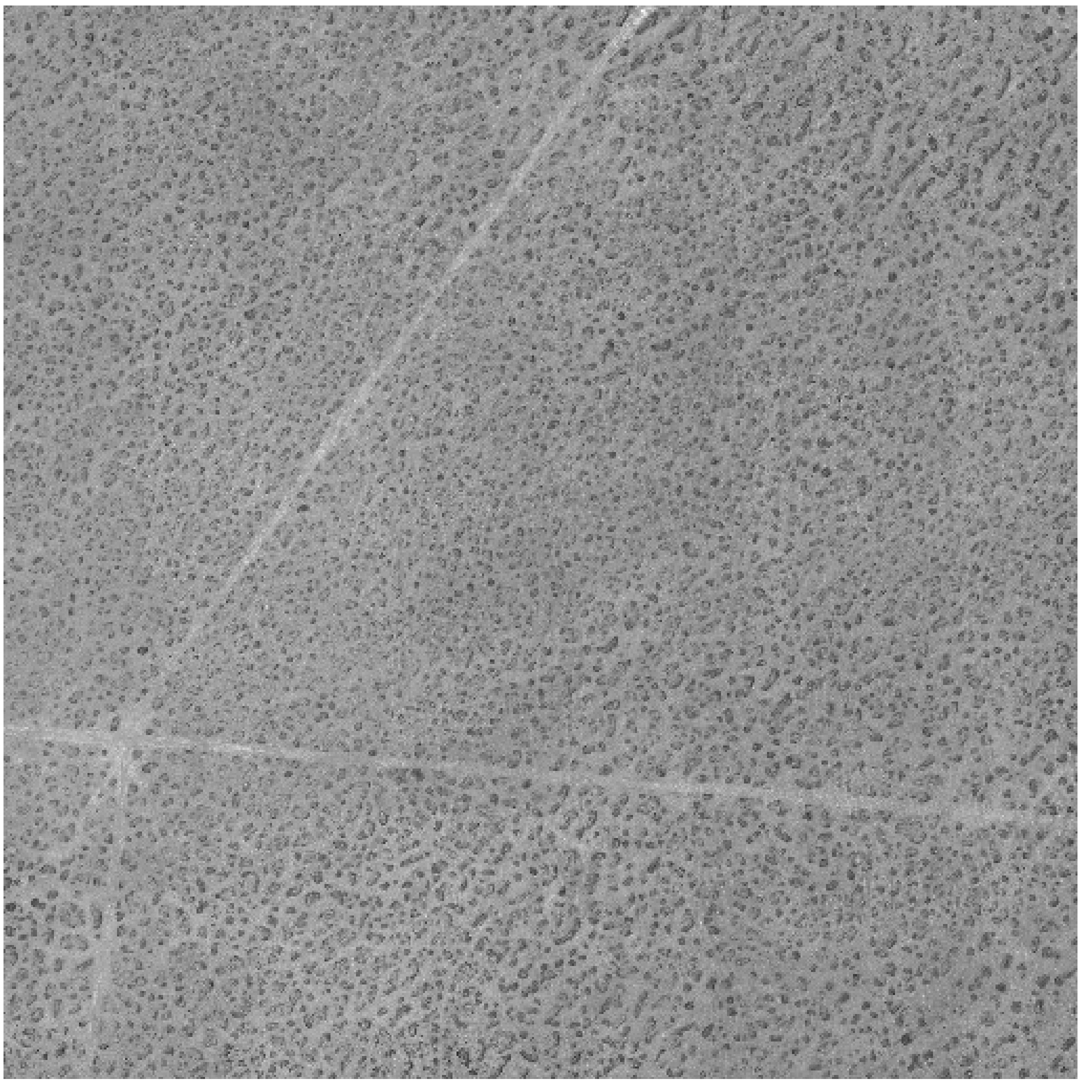


Figure 98. Intensity image (IR-albedo) for tile 384000_3572000 (1-m resolution LIDAR data). Cross roads are visible between hills with vegetation.

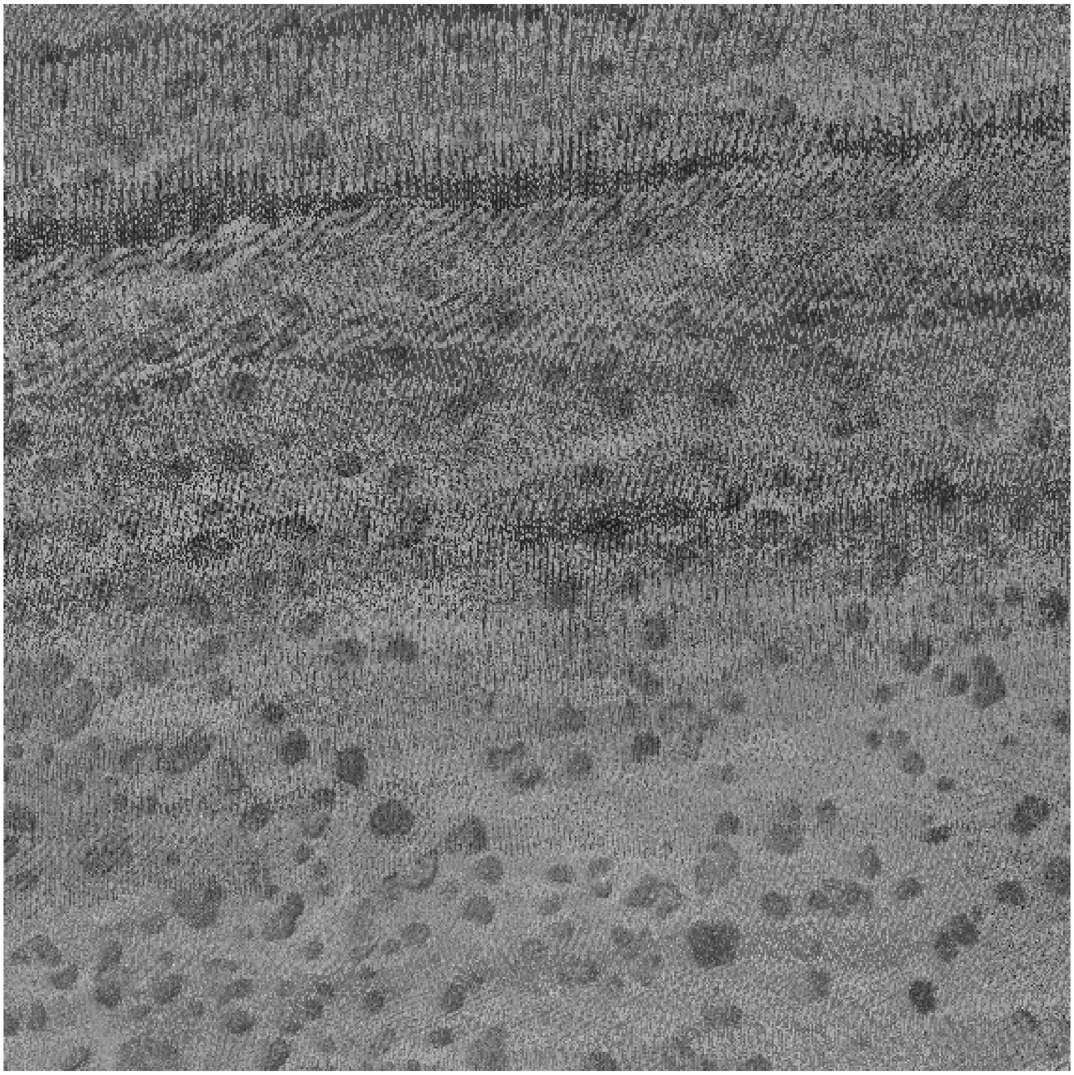


Figure 99. Intensity image (IR-albedo) for sub-tile 388000_3574341. Signatures of poor calibration of LIDAR data are visible in top part of the image. Dark spots indicate good correlation between the vegetation clusters and small hills from DTM.

With high quality intensity imagery, this information can be used as an additional signature of vegetation pixels, which are usually darker than dry soil.

Heights of Vegetation and Structures

The following data (Figures 100 and 101) shows the heights of objects (preliminary classified foliage and structures) above DTM.

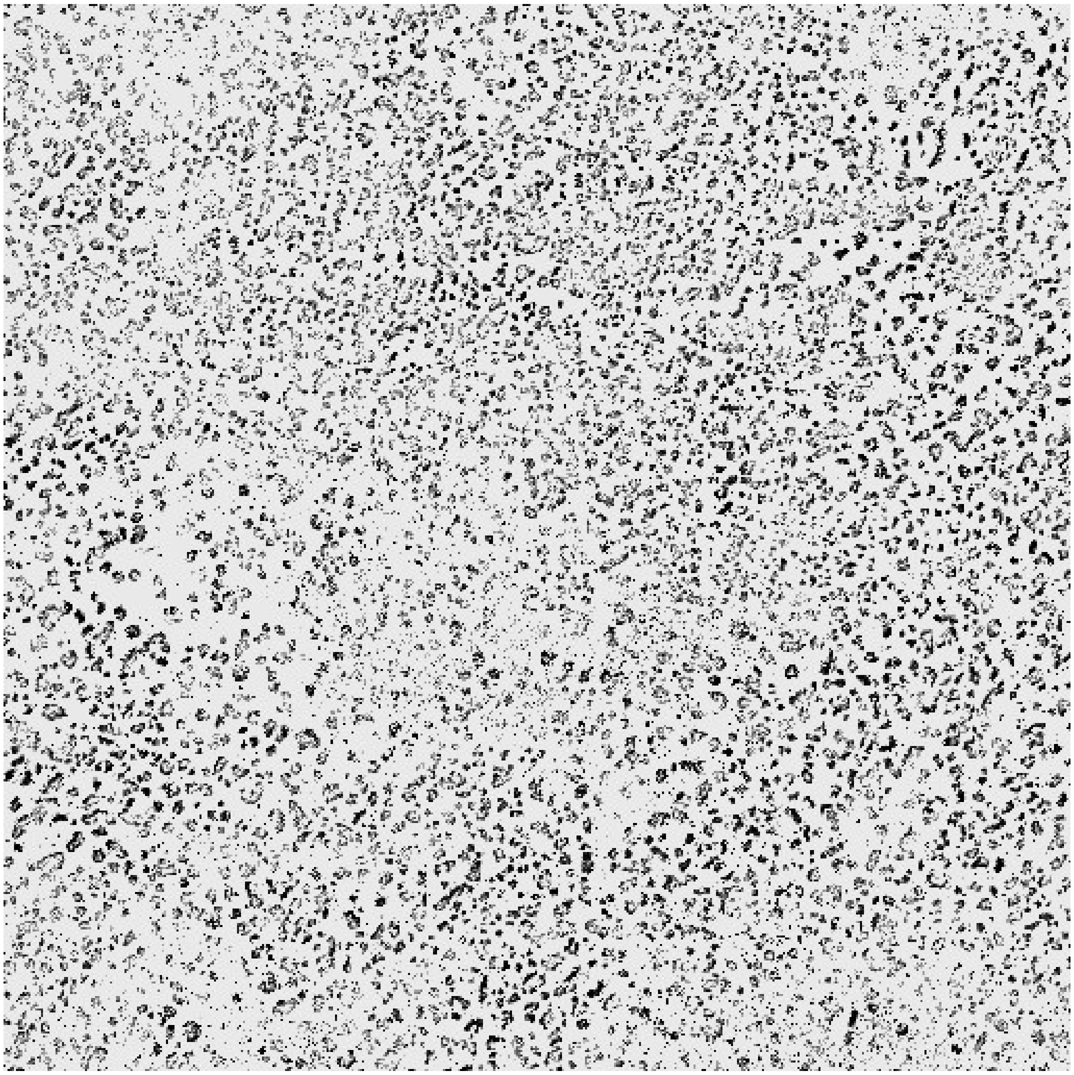


Figure 100. The distribution of pixels classified primarily as vegetation (h> 15 cm) for tile 388000_3574000 (1-m resolution LIDAR data). Each tile, sized 1,024 x 1,024 m with 1-m resolution, consists of approximately 15,000 clusters of foliage.

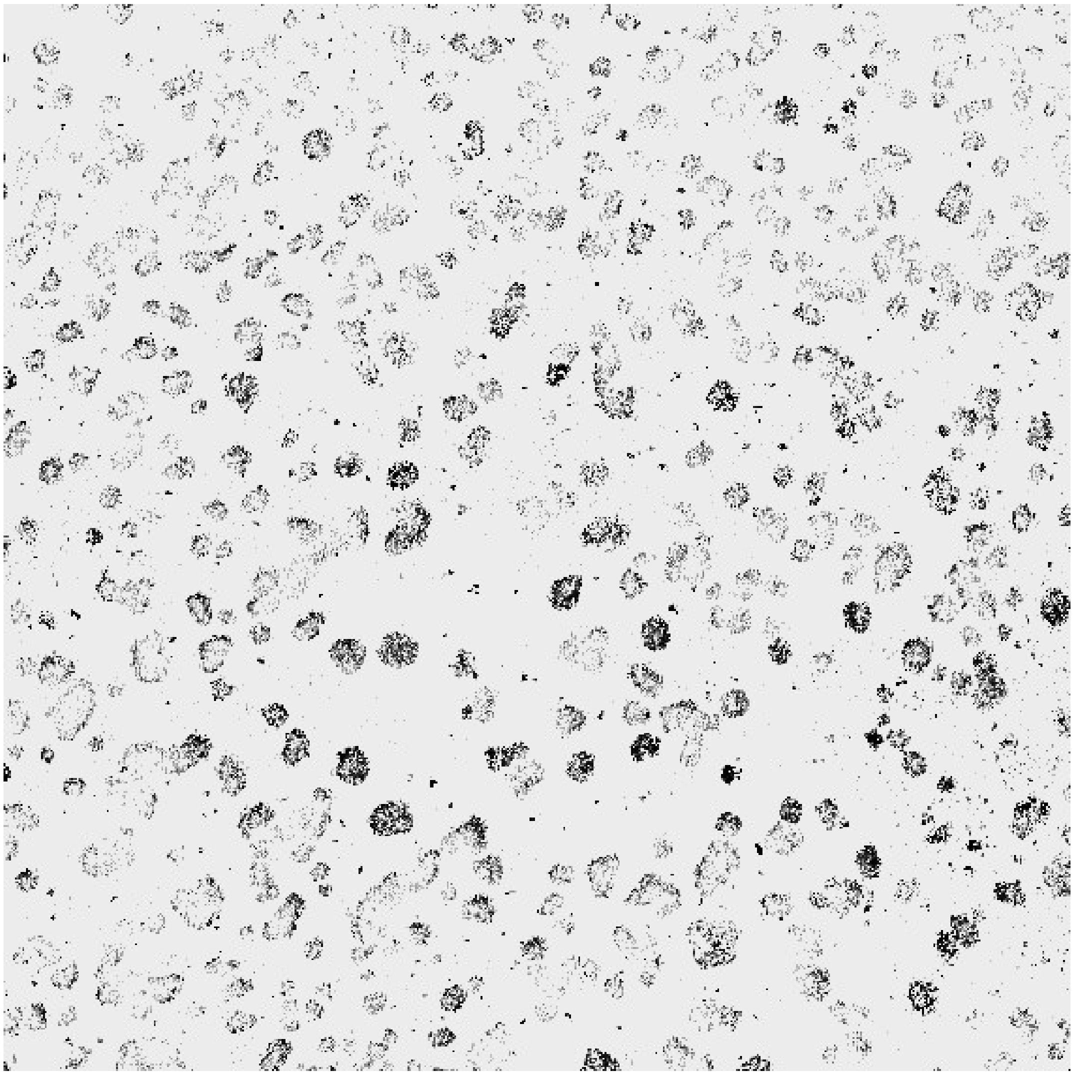


Figure 101. The distribution of pixels, classified primarily as vegetation (h> 15 cm) for tile 388000_3574341 (30-cm resolution LIDAR data). Each sub-tile, sizes 341x341-m with 30-cm resolution, consists of approximately 20,000 clusters of foliage.

Attributes or Metric File

This file contains information regarding the area of delivery; the interpolated areas in DTM and types of interpolations; and noise (Figures 102 and 103). (For details see Appendix B.) Average (for all collected data) values of coefficients for budget errors are:

	E_{pixel}	E_1	E_2	E_3	E_4	E_{noise}	
For 1-m data	0.047	0.095	0.241	0.545	1.668	0.041	(m)
For 30-cm data	0.020	0.040	0.096	0.457	1.007	0.030	(m)

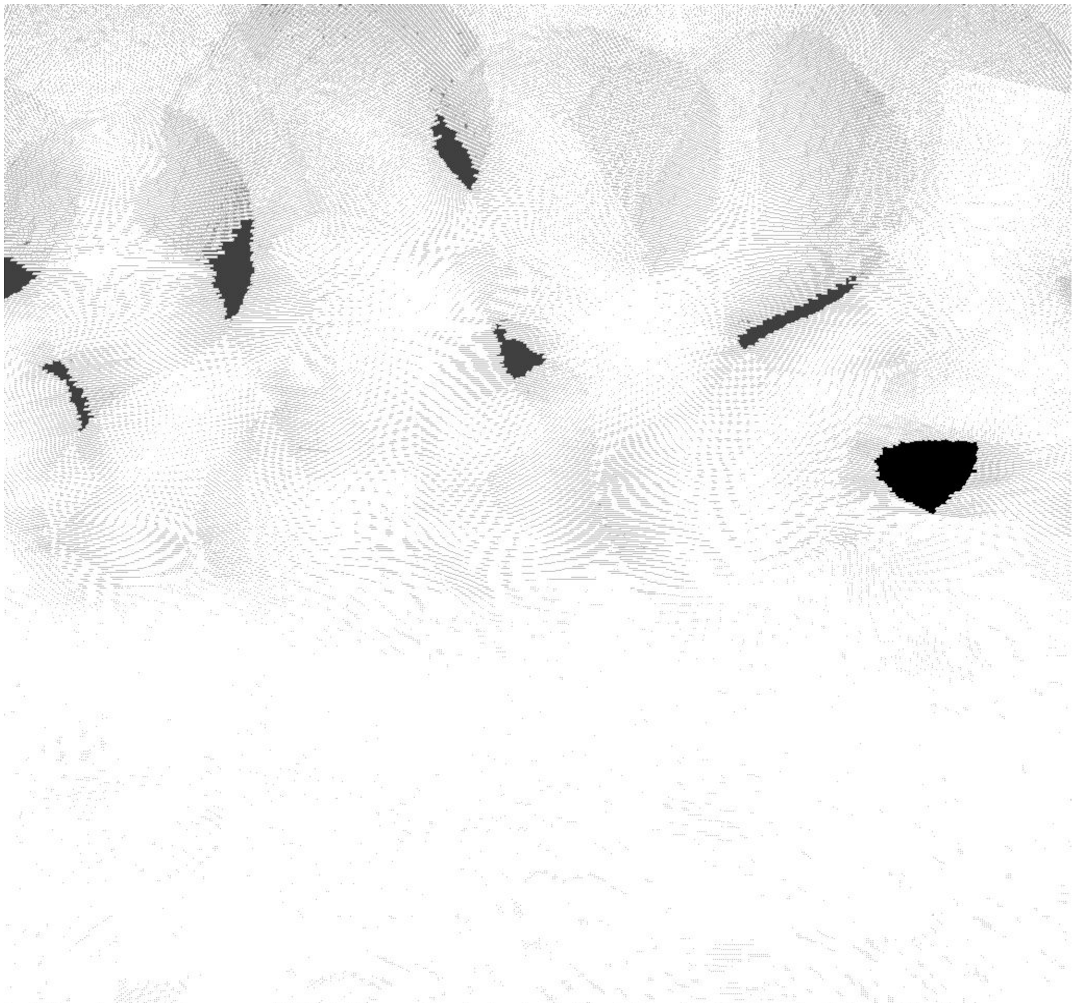


Figure 102. Imagery for Attribute or Metric File. Dark areas and points show the distribution of empty pixels in original LIDAR data for tile 381000_3582000 (1-m resolution LIDAR data). Spots in lower part of image show the distribution for a pixel of noise.

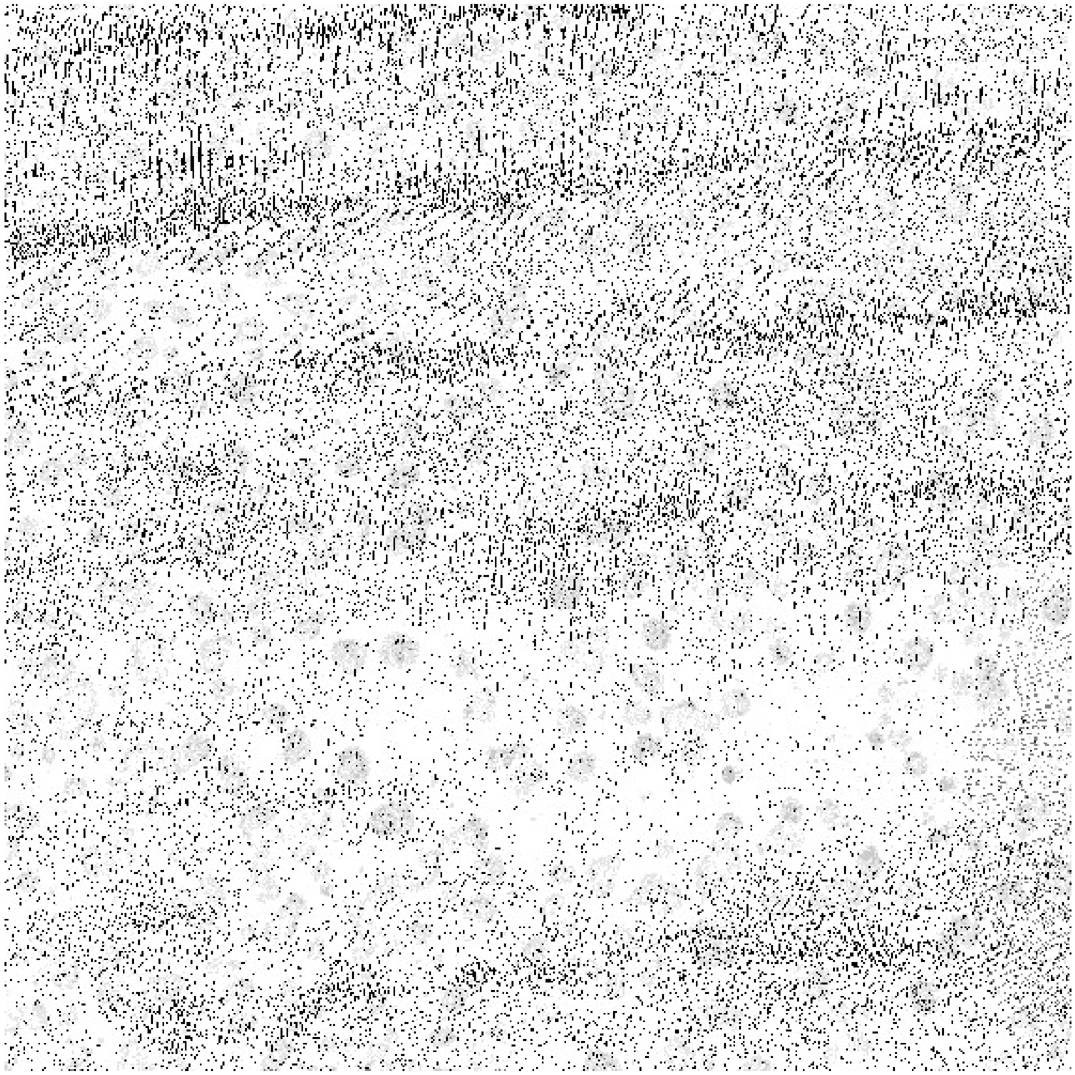


Figure 103. Imagery for attribute or metric file. Dark points show the distribution of empty pixels in original LIDAR data for tile 388000_3574341 (30-cm resolution LIDAR data). Larger and lighter spots show the distribution for a pixel of noise.

Foliage

All pixels of vegetation were sorted into clusters (a set of connected pixels) or foliage objects. The total of 1,668,667 foliage objects, from 107 tiles of data with 1-m resolution, was extracted. Average density (Figure 104) of a foliage object is approximately 20,000 per km^2 or 1 object per 50 m^2 (= 50 pixels).

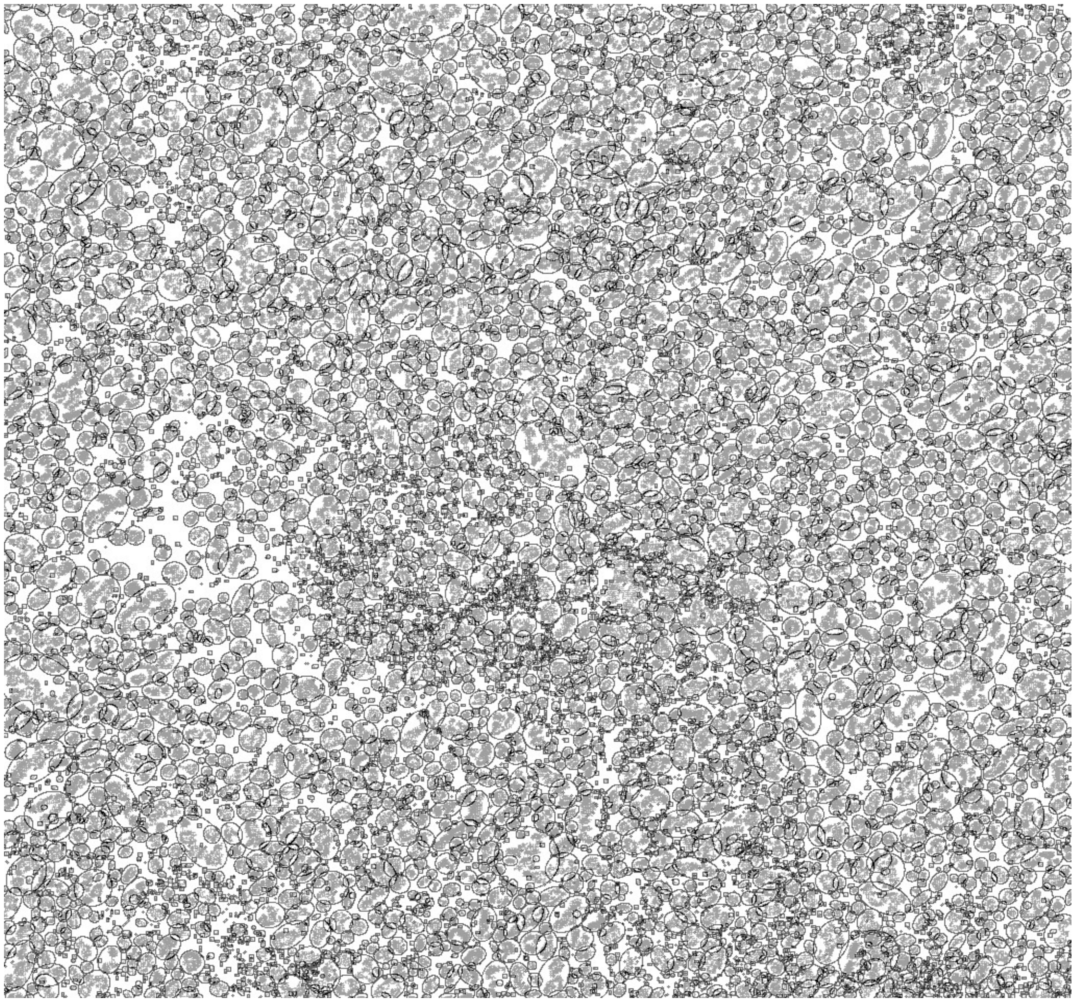


Figure 104. Imagery for foliage models, tile 388000_3574000 with 1-m resolution LIDAR data. Dark pixels show vegetation >15 cm, lighter pixels show boundary pixels with heights of 7.5–15 cm. Ellipses (darkest pixel) present integral models.

From 155 sub-tiles of data with 30-cm resolution, 3,127,057 foliage objects were extracted. Average density (Figure 105) of objects is approximately 15,000 per sub-tile or about 100,000 – 150,000 per km^2 or higher than 1 object per 10 m^2 (around 100 pixels). A comparison with Figure 104 indicates that integral ellipsoidal models for foliage, with 30-cm resolution data, are much better than for 1-m resolution data (Figure 106).

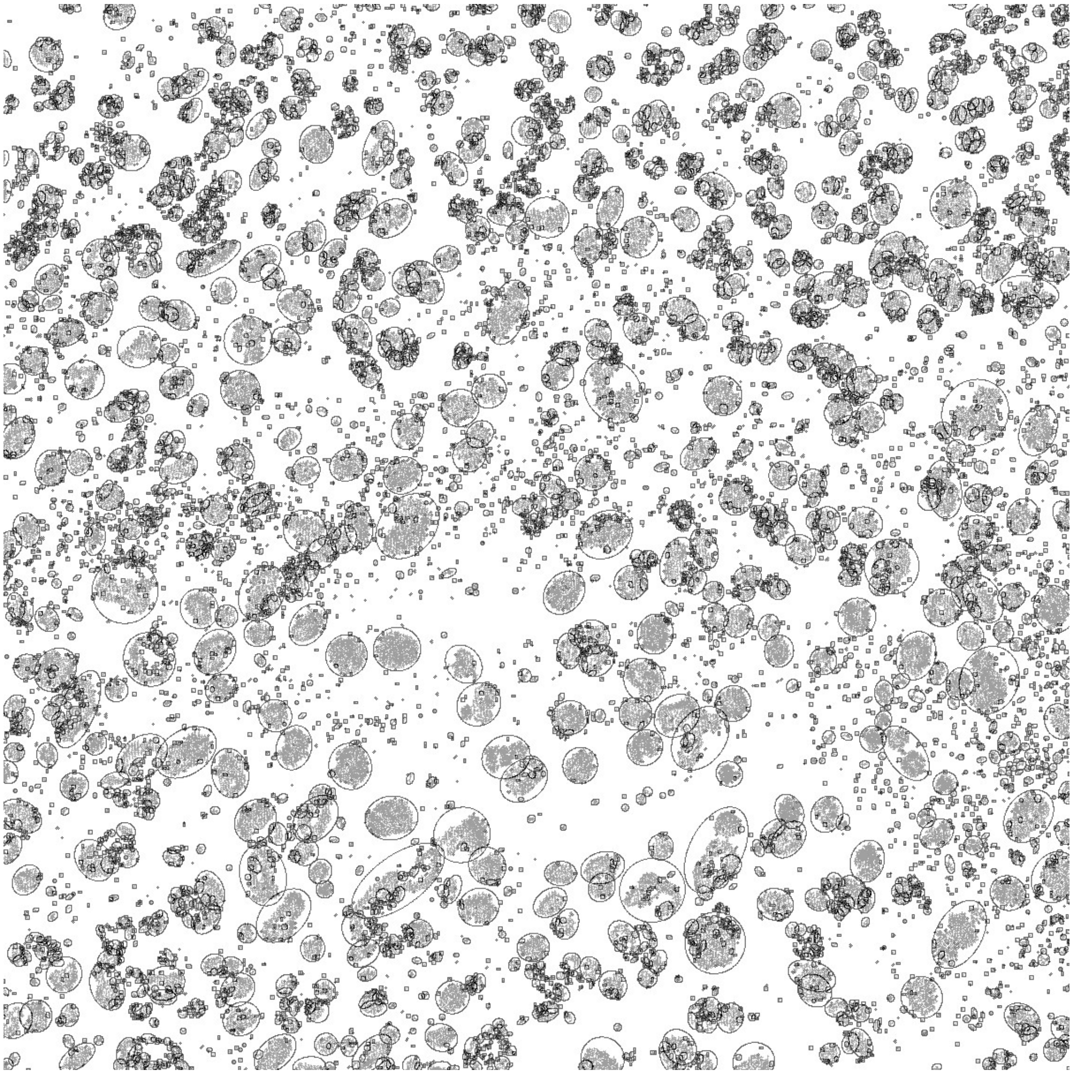
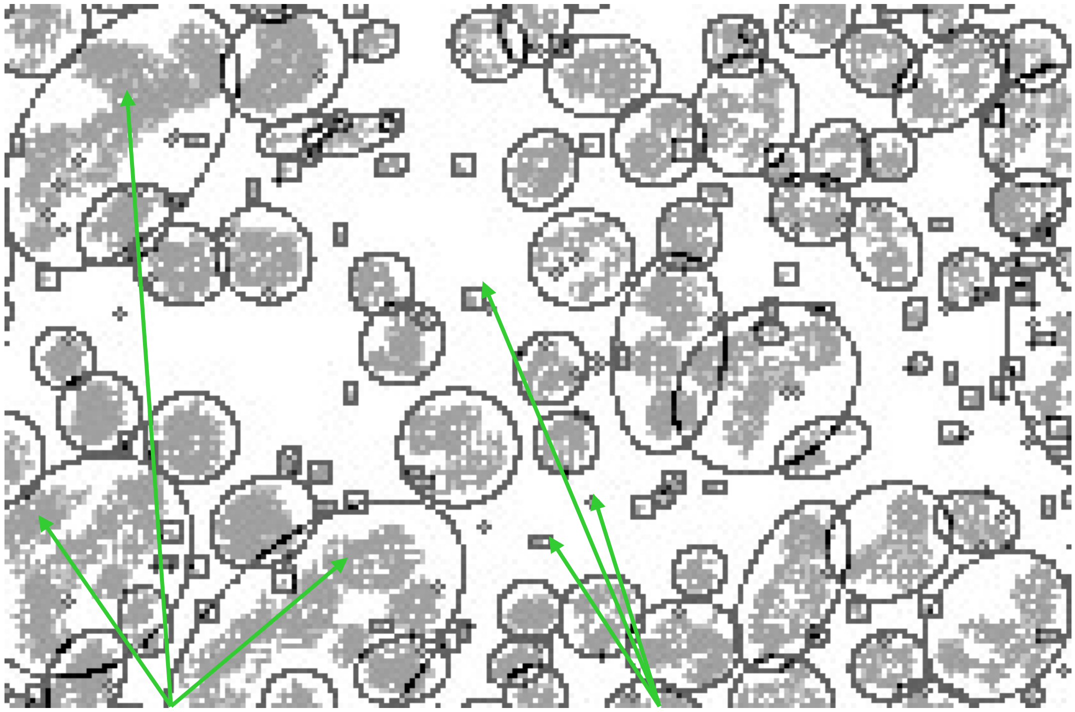


Figure 105. Imagery for foliage models from tile 388000_3574341, 30-cm resolution LIDAR data.



Large clusters of vegetation from 1-m data presented in 30-cm data (below) shows better accuracy as a set of smaller clusters.

Small clusters of 1-m data have analogs in 30-cm data (green arrows). Many small foliage objects of 30-cm data are not presented in 1-m data (blue arrows).

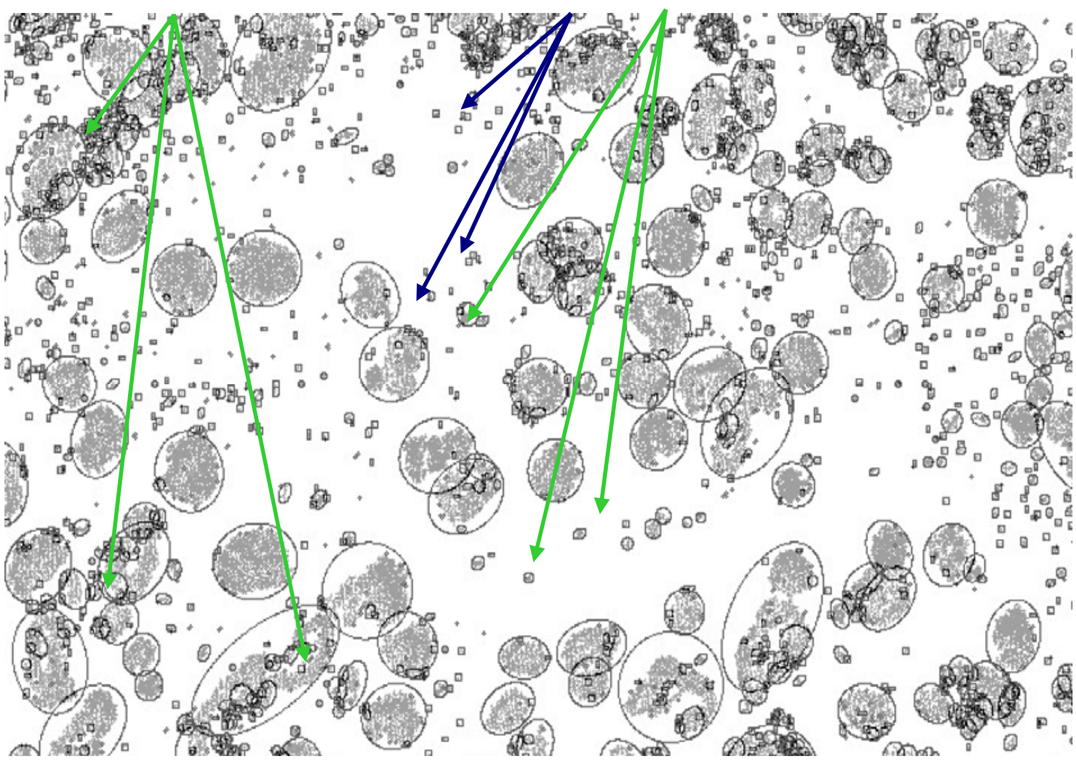


Figure 106. The comparison of vegetation models for one area and two data sets with 1-m and 30-cm resolutions using sections from tile 388000_3574000 and sub-tile 388_3574341.

Unfortunately, a large number of positive artifacts in collected LIDAR data mixed with the foliage resulting in serious damage to the vegetation models. About 50 percent of tiles with 1-m resolution have signatures of positive artifacts. The total damaged area in 1-m resolution data is estimated to be approximately 5 to 10 percent of the project area.

For 30-cm resolution data, about 70 percent of the sub-tiles were damaged by artifacts (Figure 107). The total damaged area is 30–50 percent of the collected region.

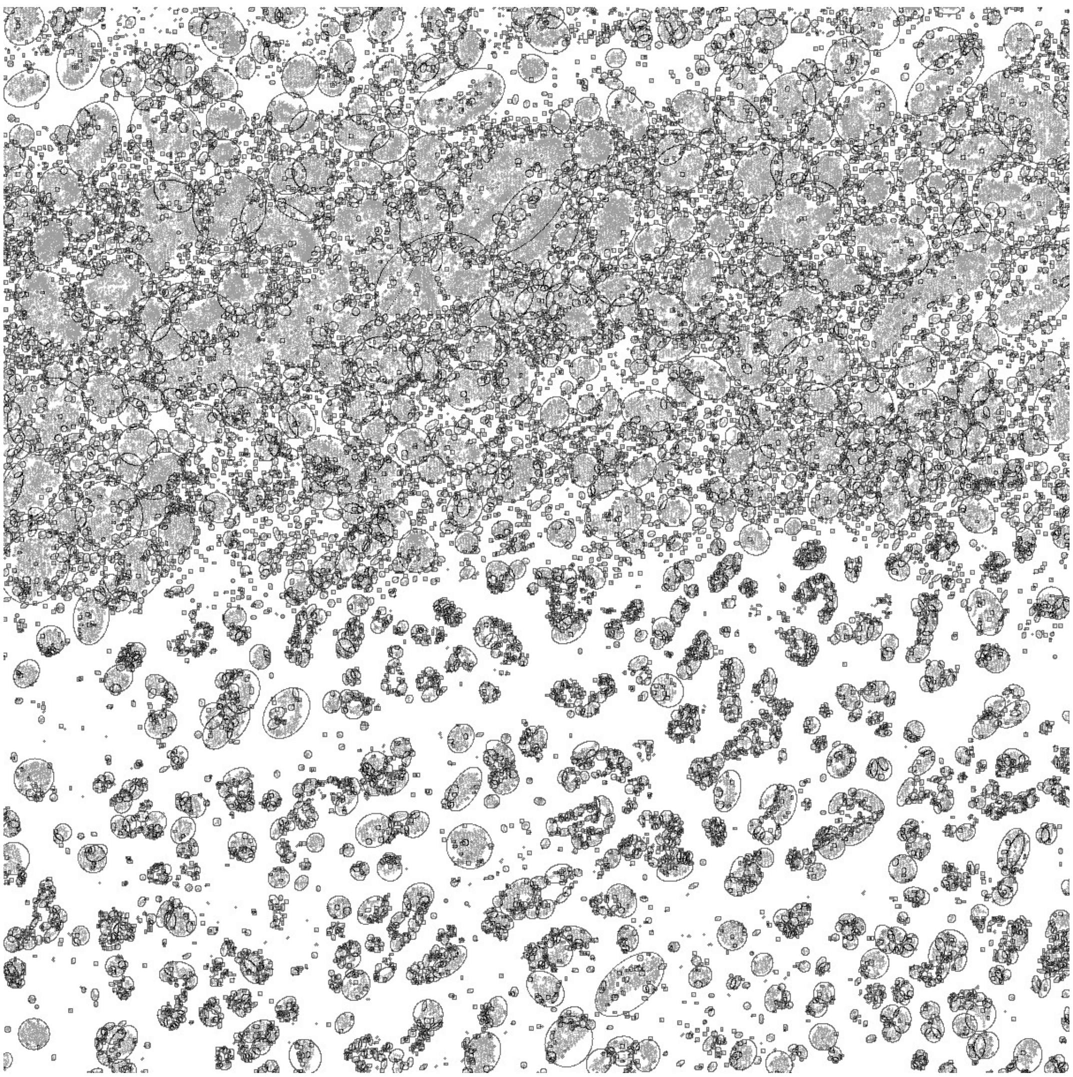


Figure 107. Foliage models in the top part of sub-tile 388000_3572682 were heavily damaged by artificial positive noise of LIDAR data. The positive noise was followed by two flight lines that tracked to other sub-tiles.

Summary and Recommendations

Summary

This project is an important step toward improving the ability to model terrains with accuracy of 30–100 cm. As result of this project, the following products were developed:

- Digital Terrain Models (bare-earth) for DTE4 COA1 on level DTED-5 (1-m resolution, for about 100 km^2), and DTED-6 (30-cm resolution, for 11–15 km^2).
- Intensity imagery that presents IR-albedo of the terrain with same resolution as DTM.
- Vegetation models in two forms: pixel level and geometrical primitives (ellipsoids); 3.1 million foliage objects were extracted from 30-cm data, and 1.7 million foliage objects were extracted from 1-m LIDAR data.
- Techniques for tracking and estimating all possible sources of errors in DTM and other layers, including budget of errors for each pixel.
- The “Description and Specification of Data Products (DTM/Foliage database)” report for this and other similar projects.
- “LIDAR and Imagery Data Collection Requirements To Support the Generation of Synthetic Scenes and Digital Terrain Models” report used the solid experience of GDATP team and lessons extracted from this project.

During the processing of LIDAR data collected from White Sands Missile Range, the team identified several serious problems in the data that affected the overall quality of the delivered data products and models including:

- Nearly 100 percent of the data was single returns, one of the primary reasons for serious underestimation of the volume of vegetation.
- The large number of positive and negative point-like and strip-like artifacts adversely affected the DTM accuracy and vegetation model.
- Poor calibration of intensity between flight lines affected the physical characterization of vegetation.

As result of the team’s additional R&D efforts, the final DTM performs well given:

- Averaging reflections in each pixel should be used, not minimal.
- Additional filtering of point-like artifacts, such as holes and holidays, can be employed by using the bare-earth DTM, which has a smooth, and predictable, surface.

These approaches work for specific case of a desert background; however, they are useless for a forested area.

The vegetation model experienced more damage by these artifacts. Vegetation modeling adhered to the following principles: "minimal damage for real objects"; and "false alarms are better than false hopes." As a result, any procedures that delete vegetation and create "false hope" for robots and vehicles (by promising free way or free access), were viewed as non-appropriate procedures. Based on this approach, one cannot use:

- averaging of the reflection's heights for vegetation will cause the loss of a large part of the bottom reflection from vegetation)
- filtering of instrumental noise from vegetation, which results in deleting part of the vegetation. This occurs because of the close geometrical parameters of local noise and small foliage.

Recommendations

For next phase of project, LIDAR data must be collected:

- According to the requirement document for LIDAR collection (prepared by GDATP/GIST team as result of this project).
- By a proven, experienced LIDAR data vendor (recommended by the GDATP/GIST team). Collection of LIDAR data from an airplane must be economical and easily available in the market.
- With independent control of quality of data immediately after delivery and before final payment (such control can be provided by GDATP/GIST team). No vendor is free from problems. Typical issues include re-processing and re-delivery original data and sometimes, re-collection of data. However, most problems can be solved using permanent control and negotiations with vendor. Providing control points and field tests is necessary for independent control.

Improvement of the vegetation models can be performed after:

- Recollection of LIDAR data for area of DTE4 COA1 during the Phase 2 project. The price of Phase 2 will be changed as percentages.
- More aggressive filtering, with permission from the customer, occurs; however, it can delete most of noise and artifacts and some part of vegetation.

The noise in different databases is not likely to be correlated with each other; however, vegetation is strongly correlated in datasets with both

resolutions. As result, most of noise in the 30-cm data can be deleted by comparing it to the 1-m data and eliminating any non-correlated data. Clearly, such an approach eliminates small foliage objects, like bunches of grass and small bushes that are presented in the 30-cm resolution data and that are not presented in the 1-m resolution data. This approach can also be used to compensate for heavily damaged parts of the sub-tiles in the 30-cm data; however, this time-consuming approach will generate non-uniformity between different parts of a sub-tile.

Future goals for improving models of terrain and vegetation include:

- A more detailed vegetation presentation (for 1-m data especially) using multiple ellipsoid models and other possible models.
- An advanced model for DTM using ellipsoidal functions.

7 Conclusions and Recommendations

This study performed an analysis of the current available IFSAR topographic data products, specifically:

1. Assessing the fidelity and quality of the commercial DSM and DTM from Intermap for the Yuma Wash area—Data Analysis
2. Developing algorithms based on AFE for LIDAR data that can be applied to processing IFSAR data—specifically to perform foliage/vegetation removal and building/structure filtering while maintaining accurate terrain profile—Virtual Surfaces Method of processing IFSAR Data
3. Assessing improvements in constructing DSM/DTM using CCS-developed methods—Results from CCS IFSAR processing.

The terrain enhancement and improvement methods described in this report rely on the use of control points obtained through ground surveys, precise LIDAR measurements, or other cadastral data, to remove uniformed and stochastic errors in IFSAR data. In denied areas, these control points are likely to be collected using LIDAR, differential GPS, or other remote sensing measurements taken by unmanned air and unmanned ground vehicles; unattended and persistent sensing; or dismounted special operations. While terrain enhancement is a function of the number of control points available, only a very small number of control points are needed to remove the large, uniformed biases in IFSAR-derived terrain models.

The following sections summarize the work done with each product, including specific conclusions and recommendations.

High Resolution Terrain Modeling Using Interferometric Synthetic Aperture Radar Data

Summary

It is clear from the data that IFSAR can be a viable technology for producing topographic data products for synthetic modeling applications. It is capable of acceptable accuracies in areas of minimal vegetation, but has limitations in areas of heavy vegetation. For this data study, two types of errors existed: the absolute accuracy, and the originally processed DTM, which consisted of an over-smoothed, and highly interpolated surface, which was not representative of the terrain. By using good control points,

the overall absolute accuracy can be improved. In this case, by shifting the data 1 m, made a significant improvement. The additional improvement can be obtained from processing the DTM from the DSM using the CCS methods as described in this report.

Conclusions

1. The Intermap DSM provides a relatively good representation of the surface; however, the DSM can be improved using the CCS techniques and V&V measurements for removing errors and biases.
2. Analysis of the Intermap DSM to the 2,481 control points of the Government Furnished Data (GFD) showed an elevation accuracy of 1.17 m RMSE.
3. Analysis of the Intermap DTM to the 2,481 GFD control points showed elevation accuracy of 1.26 m RMSE.
4. Further analysis showed that the Intermap DTM is further degraded with increase foliage density and terrain gradient or offsets.
5. Results show that the CCS AFE techniques and other terrain/surface construction methods can significantly improve the accuracy of the IFSAR derived DTM. The CCS team was able to successfully remove surface objects and foliage without adversely affecting spatial accuracy. Furthermore, the CCS methods are expected to provide substantial improvements to IFSAR derived DTMs in areas of heavy vegetation.
6. Across the set of control points, implementation of the CCS algorithms reduced the errors in the DTM from 35 to 50 percent. For 2186 control points, the Intermap DTM has an error 130 cm RMSE, and CCS DTM has an error of 65 cm RMSE.

Control points are necessary for generation of high quality DTM from IFSAR data, and a procedure for needs to be developed to optimize the V&V process.

Processing Color Imagery for White Sands Missile Randge

Summary

This project is an important step toward improving the ability to extraction of foliage objects from color imagery with accuracy of 30 cm. As result of this project, the following products were developed and delivered:

1. 353 color imagery, 300x300 m or 1000x1000 pixels. GeoTIFF. (*.tif in folder IMAGERY). Name of imagery comprising from name of

original large image and coordinates of center of pixel at left top corner of imagery with accuracy 0.1 m.

2. 323 imagery with classified foliage pixels, 300x300 m or 1000x1000 pixels. GeoTIFF.
   (vell*.tif in folder FOLIAGE IMAGERY). Total size of processed imagery is more than 29 km^2.
3. 323 files with integral models of vegetation. Total number of extracted foliage obects: 1 million and 95 thousand.
   (vint*.dat in folder INTEGRAL VEGETATION MODELS)
4. 323 files with local models of vegetation. Total number of extracted foliage obects: 1 million and 95 thousand.
   (vloc*.dat in folder LOCAL VEGETATION MODELS)
5. Polyline of roads with attributes. ArcView format; 292 roads (segments of roads) were extracted in area of 4.5 km^2 (50 imagery, 300x300 m each). Total length of extracted roads more than 42 km.
   (ArcView files for polylines of roads in folder ROADS; imagery – in folder IMAGERY)

Recommendations

1. For next phase of WSMR project, color imagery can be used for foliage extraction with quality 95 percent and level of false alarms smaller than 5 percent.
2. Discussion with customer about desired algorithm of extraction and determination of foliage objects, especially in urban area with green lawns, is necessary for improving of extraction of foliage.
3. Improvement of the vegetation models can be performed after:
   a. Improving of criteria of extraction of foliage against background with variability of brightness and spectral parameters.
   b. Improving estimation for heights of foliage using field tests and control measurements of heights of different types of vegetation.

Conclusions Regarding Verification and Validation Assessment of Interferometric Synthetic Aperture Radar

Analyzing quality and fidelity IFSAR AFE and DTM requires the use of quantitative parameters or criteria of quality AFE:

1. RMSE, which must be minimal.
2. Bias, which ideally should be minimal and close to zero.
3. Maximal percent of extracted artificial objects and trees.
4. Maximal percent of maintained details of relief.

The end results show that together these AFE algorithms and V&V method can significantly improve the Intermap DTM and provide a DTM of higher quality. The RMSE improved from over 100 cm to 65 cm. The bias was also identified and removed resulting in the improved RMSE. In all cases the data improves not only statistically, but the overall data integrity of the true surface elevations is maintained by ensuring legitimate features are extracted and the details of the relief are maintained.

Conclusions Regarding Light Detection and Ranging and Imagery Data Collection Requirements To Support the Generation of Synthetic Scenes and Digital Terrain

The most important factors that affect LIDAR data and DTM cost are:

- area of collection (both size and shape)
- density of data
- required accuracy of bare-earth data
- location of project area (proximity to airport and usable survey control points)
- level of processing to minimize artifacts (automated vs. manual processing)
- availability of ground truth checkpoints
- requirement for hydro-enforcement.

LIDAR sensors can be mounted on ground vehicles to collect 3-D surfaces around a vehicle, rather than looking downward from an aircraft. The requirements for collecting and processing ground vehicle LIDAR data are similar to the requirements for airborne LIDAR systems. The most important differences are:

- better resolution and quality due to smaller distances
- less expensive per number of shots
- increased time to cover the project area due to the lower vehicle speed
- collection area is restricted due to shadowing and access restriction.

Processing LIDAR data from ground sensor presents more challenges due to the complex 3-D geometry of collected data.

Generation of Synthetic Scenes and Digital Terrain Models Using LIDAR Data from White Sands Missile Range

Summary

This project is an important step toward improving the ability to model terrains with accuracy of 30–100 cm. As result of this project, the following products were developed:

- Digital Terrain Models (bare-earth) for DTE4 COA1 on level DTED-5 (1-m resolution, for about 100 km^2), and DTED-6 (30-cm resolution, for 11–15 km^2).
- Intensity imagery that presents IR-albedo of the terrain with same resolution as DTM.
- Vegetation models in two forms: pixel level and geometrical primitives (ellipsoids); 3.1 million foliage objects were extracted from 30-cm data, and 1.7 million foliage objects were extracted from 1-m LIDAR data.
- Techniques for tracking and estimating all possible sources of errors in DTM and other layers, including budget of errors for each pixel.
- The "Description and Specification of Data Products (DTM/Foliage database)" report for this and other similar projects.
- "LIDAR and Imagery Data Collection Requirements to Support the Generation of Synthetic Scenes and Digital Terrain Models" report used the solid experience of GDATP team and lessons extracted from this project.

During the processing of LIDAR data collected from White Sands Missile Range, the team identified several serious problems in the data that affected the overall quality of the delivered data products and models, including:

- Nearly 100 percent of the data were single returns, one of the primary reasons for serious underestimation of the volume of vegetation.
- The large number of positive and negative point-like and strip-like artifacts adversely affected the DTM accuracy and vegetation model.
- Poor calibration of intensity between flight lines affected the physical characterization of vegetation.

As result of the team's additional R&D efforts, the final DTM performs well given the following:

- Averaging reflections in each pixel should be used, not minimal.

- Additional filtering of point-like artifacts, such as holes and holidays, can be employed by using the bare-earth DTM, which has a smooth, and predictable, surface.

These approaches work for specific case of a desert background; however, they are useless for a forested area. The vegetation model experienced more damage by these artifacts. Vegetation modeling adhered to the following principles: "minimal damage for real objects"; and "false alarms is better, than false hopes." As result, any procedures that delete vegetation and create "false hope" for robots and vehicles (by promising free way or free access), were viewed as non-appropriate procedures. Based on this approach, one cannot use:

- Averaging of the reflection's heights for vegetation will cause the loss of a large part of the bottom reflection from vegetation);
- Filtering of instrumental noise from vegetation, which results in deleting part of the vegetation. This occurs because of the close geometrical parameters of local noise and small foliage.

Recommendations

For next phase of project, LIDAR data must be collected:

- According to the requirement document for LIDAR collection (prepared by GDATP/GIST team as result of this project).
- By a proven, experienced LIDAR data vendor, (recommended by the GDATP/GIST team). Collection LIDAR data from an airplane must more economical and easily available in market.
- With independent control of quality of data immediately after delivery and before final payment (such control can be provided by GDATP/GIST team). No vendor is free from problems. Typical issues include re-processing and re-delivery original data and sometimes, re-collection of data. However, most problems can be solved using permanent control and negotiations with vendor. Providing control points and field tests is necessary for independent control.

Improvement of the vegetation models can be performed after:

- Recollection of LIDAR data for area of DTE4 COA1 during the Phase 2 project. The price of Phase 2 will be changed a percentages.
- More aggressive filtering, with permission from the customer, occurs; but can delete most of noise and artifacts and some part of vegetation.

The noise in different databases is not likely correlated with each other; however, vegetation is strongly correlated in datasets with both resolutions. As result, most of noise in the 30-cm data can be deleted by comparing it to the 1-m data and eliminating any non-correlated data. Clearly, such an approach eliminates small foliage objects, like bunches of grass and small bushes that are presented in the 30-cm resolution data and not presented in the 1-m resolution data. This approach can also be used to compensate for heavily damaged parts of the sub-tiles in the 30-cm data; however, this time-consuming approach will generate non-uniformity between different parts of a sub-tile.

Future goals for improving models of terrain and vegetation include:

- a more detailed vegetation presentation (for 1-m data especially) using multiple ellipsoid models and other possible models.
- an advanced model for DTM using ellipsoidal functions.

Acronyms and Abbreviations

Term	Spellout
AFE	automated feature extraction
ASCII	American Standard Code for Information Interchange
ASPRS	American Society for Photogrammetry and Remote Sensing
ATEC	Army Test and Evaluation Command
CCS	Computational Consulting Services
CEERD	U.S. Army Corps of Engineers, Engineer Research and Development Center
COA	Common Operating Area
COTS	commercial off-the-shelf
CVA	Consolidated Vertical Accuracy
DC	direct current
DEM	digital elevation model
DSM	digital surface model
DTED	Digital Terrain Elevation Data
DTM	digital terrain model
ERDC	Engineer Research and Development Center
ERDC-CERL	Engineer Research and Development Center, Construction Engineering Research Laboratory
FCS	Future Combat System
FEMA	Federal Emergency Management Agency
FVA	Fundamental Vertical Accuracy
GDATP	General Dynamics Armament and Technical Products (team)
GFD	Government Furnished Data
GIST	Greenwich Institute for Science and Technology
GPS	global positioning system
IFSAR	Interferometric Synthetic Aperture Radar
IR	infrared
LIDAR	Light Detection and Ranging
NDEP	National Digital Elevation Program
PDOP	positional dilution of precision
QA	quality assurance
QC	quality control
R&D	research and development
RMSE	Root Mean Square Error
SVA	Supplemental Vertical Accuracy
TD	technical director
TEC	Topographic Engineering Center
TNT	trinitrotoluene
TR	Technical Report

Term	Spellout
TRAC-WSMR	U.S. Army Training and Doctrine Command (TRADOC) Analysis Center-White Sands Missile Range (TRAC-WSMR)
TRADOC	U.S. Army Training and Doctrine Command
URL	Universal Resource Locator
V&V	Verification and Validation
VDOP	Vertical Dilution of Position
VSM	Virtual Surface Method (VSM)
WSMR	White Sands Missile Range
WWW	World Wide Web
YPG	Yuma Proving Grounds

Appendix A: Ground-Based LIDAR Systems

A.1 SideSwipe (Mosaic Mapping Corporation)

One ground-based LIDAR system is SideSwipe operated by Mosaic Mapping Corporation (http://www.mosaicmapping.com).

Depending on the mission and the objects to be scanned, a tilt angle of between 70 and 110 degrees is applied to the laser that points it at the required objects. As the vehicle moves along the street or any other corridor of interest, the laser constantly scans the entire right–hand–side of its trajectory with a swath width of 60 degrees (Figure A1).

Figure A1. SideSwipe truck mounted LIDAR system.

SideSwipe's theoretical accuracy is a function of several component parts, as outlined in Table A1.

Table A1. SideSwipe error sources.

Scanning Distance (m)	10	50	100
GPS Accuracy (m) [1]	± (0.02 + 5ppm)	± (0.02 + 5ppm)	± (0.02 + 5ppm)
Beam Divergence (m)	0.030	0.150	0.300
IMU Accuracy (0.05 deg)	0.009	0.044	0.089
Laser Distance Accuracy (m) [2]	0.020	0.020	0.020
Notes: [1] Dual–frequency double differenced data, equipment operated in kinematic mode. [2] Assumes reflections off hard surfaces.			

The data in Table A1 indicate that the system will achieve horizontal accuracies of the order of 4 or 5 cm regardless of scanning distance, and vertical accuracies of the order of 5 or 6 cm at a scanning range of 10 m. The vertical accuracy degrades with increased scanning distance. Therefore, SideSwipe essentially turns LIDAR on its side, changing vertical to horizontal and horizontal to vertical. Similar accuracies are achieved in both the horizontal and vertical components. SideSwipe's operating costs are minor compared with those of the helicopter system upon which it is derived. The helicopter system can be modified for ground use in half a day.

The graphic shown in Figure A2 is generated entirely from SideSwipe data collected at the ground level. The black areas represent areas of shadowing. The impact of shadowing can be minimized by adopting appropriate field techniques.

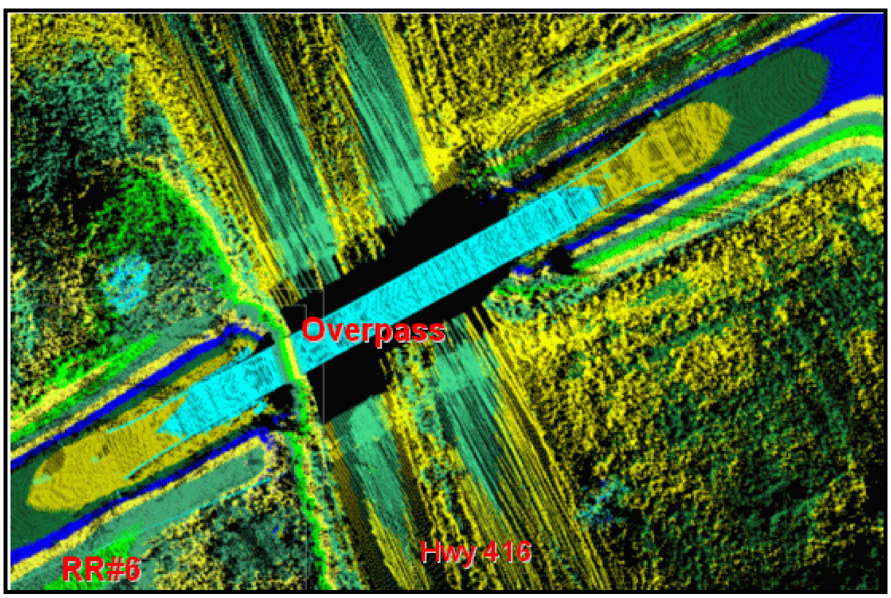


Figure A2. Digitally enhanced SideSwipe data.

A.2 ILRIS-3D Intelligent Laser Ranging and Imaging System (Optech Incorporated)

Optech's ILRIS-3D is a portable laser-based imaging and digitizing system that offers many enhanced features; including an integrated 6 mega pixel digital camera and automated true color point clouds (Figures A3 to A6). ILRIS-3D has a dynamic scanning range of 3 m to beyond 1,000 m, a minimum scan step size of 0.0014915 degrees and Class 1 laser eye safety designation. Product features include the following:

Range:

350 m (4% albedo target)

800 m (20% albedo target)

1,500 m (80% albedo target)

Accuracy (depth resolution):

3 mm for 100 m

Laser Spot Size:

D=0.17R+12, where D = diameter of spot (mm); R=Range to target (m)

D=29 mm for 100 m

Minimum Spot Spacing:

S=0.026R, where S = Spacing (mm); R=Range to target (m)

S=2.6 mm for 100 m

Data sample rate:

2,000 points/sec. Each LIDAR point also has real color from optical camera.

Figure A3. General view of portable ILRIS-3D.

The ILRIS-36D will orient the scanner system anywhere within a complete 360° radius.

Figure A4. Tokyo intensity of LIDAR returns for ILRIS-3D.

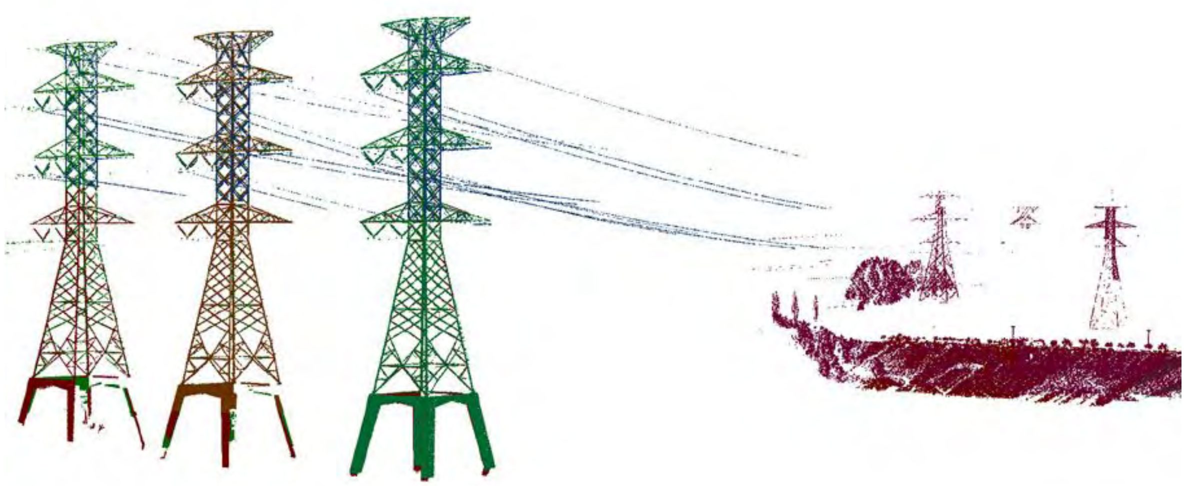


Figure A5. Colored by distance LIDAR data for power lines from ILRIS-3D.

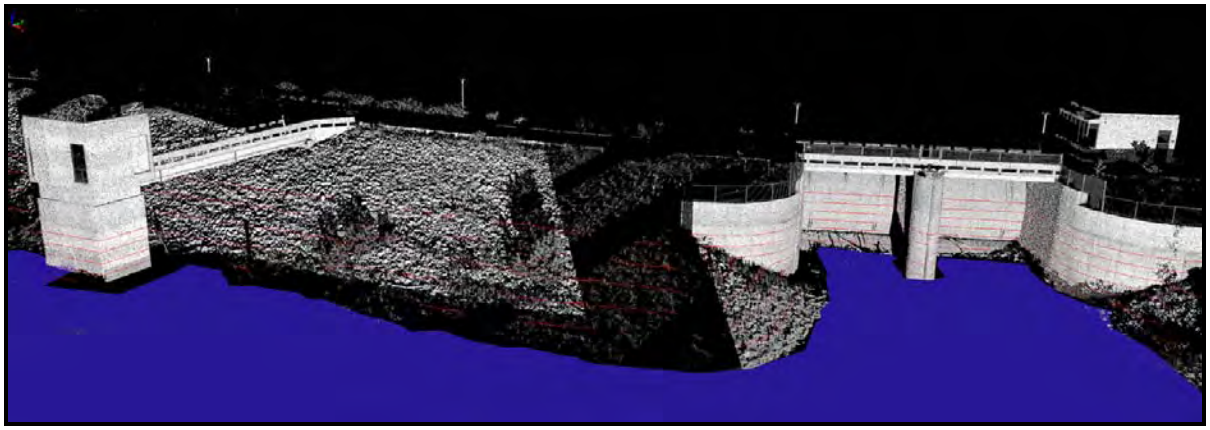


Figure A6. LIDAR point cloud for dam and ground.

Appendix B: Description and Specification of Data Products (DTM/Foliage/Surface Objects Database)

B.1 General Description of Data Products for White Sands Project

The size of the tiles for White Sands Project was:

- for 1-m resolution data – 1,024 x 1,024 m with overlap of 24 m on right and top sides
- for 30-cm resolution data – 341 x 341 m (large 1,024x1,024 m tile was split into 9 sub-tiles without overlap,

B.1.1 DTM GeoTIFF Format

Each pixel of the DTM includes an absolute height of bare earth without vegetation, structures and noise. The gridded DTM have interpolated parts and areas due to:

- missing original LIDAR data
- fixing artifacts (pits or underground reflections)
- extracting objects (foliage, structures, noise, etc.),

Information regarding the interpolated areas and types of interpolation will be collected in Attributes or Metric File (see Sections B2 and B5), which serves as a confidence metric for each step of DTM generation.

B.1.2 Intensity Imagery: Gridded File in GeoTIFF Format

Each pixel of imagery includes intensity of LIDAR returns or IR (~1 micron) albedo. Intensity measures the percentage of the reflectance returned from the output laser pulse. Interpolation of empty areas and holes in intensity imagery follows the same rules as interpolation of empty areas of DTM.

B.1.3 Heights of Vegetation and Structures: Gridded File in GeoTIFF Format

Each non-zero pixel of this file consists of the relative heights of all objects higher that DTM + noise (DTM+15 cm for White Sands project). Empty areas and gaps in original data cannot be classified as vegetation or struc-

tures in this file. Determination of vegetation in this file follows a rule for relative heights >15 cm and different from determination of vegetation in foliage the database (Section B5). Foliage database also includes boundary pixels with h > 7.5 cm contacted with pixels of high foliage h > 15 cm.

B.1.4 Attributes or Metric File in GeoTIFF format (gridded file)

This file keeps information regarding:

- area of delivery
- interpolated areas in DTM and types of interpolations
- noise (see detailed description below),

This file can be converted to ASCII file, analyzed and separated into different layers/files of information about noise only, or attributes, or Z-accuracy of interpolated areas etc.

B.1.5 Foliage Database in ASCII Format

This file is a list of objects of vegetation. The description of each object of vegetation has two levels:

- pixel level (list of pixels of an object with all available information for each pixel)
- model level with parameters of model and integral parameters,

Detailed description of foliage database is below. Specification for these data products is also applicable to for other information or future needs. For example, any other imagery (optical or infrared imagery) can follow the format of GeoTIFF/gridded deliverables in Sections B1–B1.3 (with metric file like the deliverable in Section B1.4). Files for other objects (buildings, water, roads, cars etc) can follow the format of ASCII-type deliverable in Section B1.5.

B.2 Description of Attributes or Metric File

Each pixel of this GeoTIFF file consists of a float number. For example

- 0.11, or
- 50000.05, or
- 62000.0

B.2.1 Legend for 5th number (tens of thousand): *"Missing and interpolated original data inside and outside of delivery area."*

0 – data exist inside of delivery area.

This attribute value indicates a **maximal** level of confidence (~ errors original LIDAR data plus error of gridding – see E_{pixel} in Section B4).

1. Missing data inside of delivery area. This attribute value indicates interpolation by type #1: data for the pixel was interpolated from nearest pixels (at least one of the nearest 8 pixels has data). This attribute indicates a **high** level of confidence for interpolated height; it depends on the type of relief and sloppiness of the surfaces.
2. Missing data inside of delivery area. This attribute value means interpolation by type #2 (curve filling) that is applicable to two typical cases:
   a. Area of the gap/hole (after interpolation by type #1) smaller than or equal to 25 pixels (height for pixels of the hole was interpolated from pixels with data on boundary of the hole)
   b. Length of interpolation (distance to pixel with data) is 4–5 pixels or smaller.

   This value of attribute indicates a **middle** level of confidence for interpolated height.
3. Missing data inside of delivery area. The interpolation identified by #3 was flat filled using the average height of the pixels on the boundary of the empty area. The area of the hole (after interpolation by type #1) is larger than 25 pixels and smaller than or equal to 2,500 pixels. This attribute value indicates a **low** level of confidence for interpolated height.
4. Missing data inside of delivery area. The interpolation marked by type #4: data was flat filled using the minimal value of the height of the pixels on the boundary of the empty area. The area of the hole (after interpolation by type #1) is larger than 2,500 pixels. This attribute value indicates a **minimal** level of confidence for interpolated height. For a flat area like White Sands, the error for interpolating the heights of holes in the region of 0.5 km is approximately 1–2 m.
5. Data exist outside of delivery area (other information similar to the attribute **0**).
6. Missing data (interpolated by type #1, see the attribute **1**) outside of delivery area.
7. Missing data (interpolated by type #2, see the attribute **2**) outside of delivery area.
8. Missing data (interpolated by type #3, see the attribute **3**) outside of delivery area.
9. Missing data (interpolated by type #4, see the attribute **4**) outside of delivery area.

Delivery area for 1-m resolution data is an area 10 x10 km. Other northern tiles are considered "outside of delivery area."

All 30-cm resolution data within the 10 km x 10 km is considered "inside the delivery area."

Quantitative parameters for **maximal /high /middle /low /minimal** levels of confidence for interpolated heights will vary for different data and types of geographic areas. Accurate determination of such parameters is possible using a large enough set of control points.

For the White Sands project, levels of confidence that indicate the magnitude of errors can be established as:

Levels of confidence:	maximal	high	middle	low	minimal
1-m resolution:	20 cm	30 cm	50 cm	80 cm	200 cm
30-cm resolution:	15 cm	20 cm	30 cm	65 cm	150 cm

B.2.2 Legend for 4th number (thousands): *"Attributes of LIDAR data during processing."*

0 – no change of data during processing.

1 – data was interpolated after deleting foliage or structures. Interpolation by type #1.

2 – data was interpolated after deleting foliage or structures. Interpolation by type #2.

3 – data was interpolated after deleting foliage or structures. Interpolation by type #3.

4 – data was interpolated after deleting foliage or structures. Interpolation by type #4.

5 – data was interpolated due to artifacts in original data (pits or abnormal low reflections). If data was interpolated with the attribute **#5**, the attribute of this pixel cannot be changed to attributes **#1–4** during next steps of processing.

6 – data not usable.

7 – conflicting data.

8 – reserved for additional information. Interpolated by type #3.

9 – reserved for additional information. Interpolated by type #4.

For the White Sands project, only attributes from zero to 5 were used.

B.2.3 Legend for numbers No. 2–3 (between 10 and 999.99)

Reserved for other possible attributes.

B.2.4 Legend for number #1 and fractional numbers (smaller that 10): *"Height of noise and artifacts."*

All heights above ground (DTM), smaller than 10 m, represent information about noise (0<h<15 cm for the White Sands project).

Noise pixels cannot be overlapped with interpolated pixels (without original LIDAR data).

Noise pixels can be overlapped with attribute 50,000 ("outside of area delivery") only.

B.3 Samples of Data for Attributes or Metric file

B.3.1 0.11

- data exist inside of delivery area
- data do not change during processing
- noise with height 0.11 m.

B.3.2 50000.05

- data exist outside of delivery area
- data do not change during processing
- noise with height 0.05 m.

B.3.3 62000.0

- missing data outside of delivery area
- data was interpolated by type #1 during fixing holes in original LIDAR data
- **high** level of confidence for interpolated height
- data was interpolated by type #2 during processing and foliage/structure deleting
- **middle** level of confidence for interpolated height
- no noise.

B.4 Budget of Errors Using Metric File

Total error due to different sources of errors and several steps of interpolation can be estimated by the following equation:

$$Z_{real\_ground} = Z_{DTM} \pm E_{LIDAR} \pm E_{pixel} \pm E_1 \pm E_2 \pm E_3 + E_4 + E_{noise} \quad (1)$$

where:

Z_{DTM} is the deliverable I.1.

E_{LIDAR} is error of original LIDAR data (for White Sands project probably ~ 10–15 cm for 1-m resolution data and ~5–10 cm for 30-cm resolution data)

E_{pixel} is error due to sorting of LIDAR data into a grid where each pixel has only one level of height. (This does not correctly present the slope surface). Typical samples of different error terms for the White Sands project include the following:

E_1 is error due to interpolation #1.

E_2 is error due to interpolation #2.

E_3 is error due to interpolation #3.

E_4 is error due to interpolation #4. (Sign is plus only due to type of filling gaps by minimal value of heights from boundary pixels)

E_{noise} is positive error due to deleting noise from DTM. Maximal E_{noise} is 15 cm. Average E_{noise} ~ 4 cm (see below).

Usually many terms of this equation are zero (pixels do not have all 4 types of interpolations).

For pixels without any interpolations (flags/attributes = 0 or 50,000), use the simple equation:

$$Z_{real\_ground} = Z_{DTM} \pm E_{LIDAR} \pm E_{pixel} + E_{noise} \quad (2)$$

or with numbers (estimation for the White Sands project without using control points):

$$Z_{real\_ground} = Z_{DTM} \pm (10\text{-}15)\text{ cm} \pm 5\text{ cm} + 4\text{ cm} \quad (3)$$

Different terms of error for each processed tile can be found in file **coefficients_for_errors.dat.**

Typical samples from file coefficients_for_errors.dat are:

- For 1-m resolution (in m):

coef_388e3574nwgs84_raw_388000_3574000.dat

E_{pixel}	E_1	E_2	E_3	E_4	E_{noise}
0.049	0.099	0.254	0.586	1.687	0.038

- For 30-cm resolution (in m):

coef_388e3574nwgs84_raw_hr_388000_3574341.dat

0.020	0.039	0.093	0.482	1.247	0.037

B.5 Foliage Database

Each foliage object in this database is a cluster (a set of pixels, connected by sides, not by corners) from two possible types of pixels:

1. Pixels identified as vegetation, higher than 15 cm for the White Sands project.
2. Pixels identified as noise, connected by sides and corners with a height between 7.5 and 15 cm, are classified as small vegetation.

B.6 File of Detailed Local Data (List of Pixels)

This ASCII-file consists of a list of pixels with parameters for each foliage object/cluster. Below is a file for real structures/clusters No. 291-294, 1-m resolution data:

X	Y	Ncluster	hmax(m)	hmid(m)	hmin(m)	HDTM(m)	Attribute	Intensity	Class
579	31	291	0.45	0.22	0.01	1235.43	1000.00	34.30	3
580	32	291	0.40	0.17	-0.19	1235.35	1000.00	32.00	3
581	31	291	0.38	0.31	0.21	1235.78	2000.00	30.90	3
583	29	291	0.12	0.02	-0.03	1235.65	0.12	35.50	2
578	31	291	0.29	0.06	-0.03	1235.26	1000.00	33.70	3
581	32	291	0.22	0.17	0.10	1235.46	1000.00	33.90	3
582	31	291	0.23	0.16	0.06	1235.68	1000.00	32.20	3
583	28	291	0.08	0.00	-0.08	1235.55	0.08	40.30	2
583	30	291	0.15	0.08	0.01	1235.68	1000.00	38.60	3
578	32	291	0.13	0.00	-0.14	1235.04	0.13	33.90	2
581	33	291	0.16	0.00	-0.15	1234.99	1000.00	38.00	3
582	32	291	0.20	0.02	-0.11	1235.36	1000.00	33.40	3
583	31	291	0.29	0.13	0.00	1235.57	1000.00	34.60	3
627	21	292	0.15	0.02	-0.10	1235.45	1000.00	37.00	3
633	21	293	0.08	0.00	-0.15	1235.54	0.08	35.80	2

X	Y	Ncluster	hmax(m)	hmid(m)	hmin(m)	HDTM(m)	Attribute	Intensity	Class
633	22	293	0.13	0.06	-0.02	1235.64	0.13	35.20	2
634	22	293	0.21	0.07	-0.03	1235.51	1000.00	36.30	3
671	21	294	0.28	0.11	-0.06	1236.34	1000.00	33.50	3
671	22	294	0.38	0.33	0.24	1236.42	1000.00	29.40	3

B.6.1 Legend

X, Y – coordinate of each pixel of cluster (object) of vegetation (in pixels).

$N_{cluster}$ – number of cluster or structure on the tile (each tile has new numbering). In White Sands each tile (1x1 km) with 1-m resolution has 10–20 thousand objects/clusters of vegetation.

h_{max} – maximal relative height of vegetation above DTM.

h_{mid} – average relative height of vegetation above DTM (calculated as average heights of all LIDAR returns in the pixel) or height of optical depth.

h_{min} – minimal relative height of LIDAR return above DTM in this pixel. Positive value of h_{min} means that LIDAR shots do not reach the ground. A negative value indicates that the minimal height from all shots is smaller than the DTM calculated from the surface of the average heights in pixels. In most cases, a negative value is typical for a pixel on the slope surface, when minimal heights are 1–2 ft less than average heights.

These parameters (**h_{max}, h_{mid}, h_{min}**) provide the following important information:

a. Relative density of top of foliage in the pixel:

$$Top\_density = (h_{mid} - h_{min}) / (h_{max} - h_{min})$$

This parameter also presents the foliage structure in this pixel.

Top_density ~ 0 (or $h_{mid} \sim h_{min}$) means that the most dense part of foliage in this pixel is close to the ground.

Top_density ~ 1 (or $h_{mid} \sim h_{max}$) means that most dense part of foliage is close to the top of the foliage.

Top_density ~ 0.5 presents middle case or uniform distribution of foliage density in a pixel.

b. Relative density of the bottom part of foliage in the pixel

Bottom_density = $(h_{max} - h_{mid}) / (h_{max} - h_{min})$

Together Top_density + Bottom_density =1

If h_{min} <0, than the Bottom_density can only apply to heights above 0, above ground height.

If h_{min} >0, it is then possible to consider a region of heights, from height of ground (H_{DTM}) to $H_{DTM}+h_{min}$ as completely non-transparent part of foliage or sand mounds or other similar (non-transparent) objects deleted from terrain.

H_{DTM} – height of DTM in the pixel. Absolute maximal height of vegetation is $H_{veg}=\mathbf{H_{DTM}+h_{max}}$

Intensity – (in %) intensity of reflection of laser beam or IR-albedo (~1 micron).

Attribute – information from attribute/metric file about this pixel (see Section B2).

Class of vegetation:

3 – high vegetation (pixel from vegetation file with h_{max} >15 cm)

2 – low vegetation (pixel from noise file with h_{max} = 7.5–15 cm)

File of Models and Integral Data

This ASCII-file consist list of objects (clusters of foliage) and parameters for each object:

Sample from 1-m resolution data:

$N_{cluster}$	Line	X_c	Y_c	Z_c(m)	H_{aver}(m)	H_{max}(m)	Structure	I_{aver}	R_{max}	Area(pixels)
8609	1	201	575	1240.55	0.33	0.91	0.52	26.59	4.47	44
$N_{cluster}$	Line	X_{el}	Y_{el}	Z_g(m)	Z_{el}(m)	a	b	c_{max}	c_{ave}	Angle(deg)
8609	2	201	575	1240.55	1239.64	5.39	5.39	1.66	0.92	201.80
8610	1	263	572	1239.55	0.09	0.09	0.53	32.50	0.00	1
8610	2	263	572	1239.55	1239.55	1.00	1.00	0.09	0.09	0.00

Sample from 30-cm resolution data:

9592	1	758	199	1247.97	0.21	0.66	0.61	28.02	12.04	97
9592	2	760	199	1247.70	1247.11	13.81	7.74	2.82	1.11	231.34
9593	1	770	190	1246.36	0.23	0.23	0.41	24.70	0.00	1
9593	2	770	190	1246.36	1246.36	1.00	1.00	0.23	0.23	0.00

Legend:

Line: each line presents a new set of integral parameters of cluster of vegetation and a model of cluster. Total number of lines of parameters for the object (cluster of vegetation) depends on the model and types of data used (parameters included multispectral, radar and other data).

Ncluster – number of clusters or structures on a tile (each tile has new numbering)

Xc, Yc – coordinates of center of mass of cluster (all X, Y – in pixels)

Zc – height of DTM in center of mass of cluster

Haver – average value of maximal relative (to DTM) heights of all pixels of vegetation in this cluster.

Hmax – maximum value of maximal heights of all pixels of vegetation in this cluster.

Structure – average value of parameter (hmid - hmin)/(hmax - hmin), which presents a foliage structure.

Iaver – average value of intensities of all pixels of cluster

Rmax – maximum distance between the center of mass of cluster and most distant pixel of cluster (in pixels)

Area – area of a cluster (in pixels)

Xel,Yel – coordinates of geometrical center of cluster (average coordinates from maximal and minimal X and Y for pixels of cluster). These coordinates will be used as ellipsoid coordinates for modeling this cluster.

Zg – height of DTM in geometrical center of cluster, or center of ellipsoid

Zel – absolute height of center of an ellipsoid = minimal height from heights of all pixels of vegetation in this cluster.

a – semi-major axis of an ellipsoid (in horizontal plane, in pixels)

b – semi-minor axis of an ellipsoid (in horizontal plane, in pixels)

cmax – vertical axis of ellipsoid that describes maximal vertical extension of cluster, in pixels

caver – vertical axis of ellipsoid that describes average vertical extension of cluster, in pixels

Angle – angle of inclination of semi-major axis of ellipsoid to coordinates axis (degrees). Angle =0 for X-axis and increase to counterclockwise direction

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
30-09-2007	Final	

4. TITLE AND SUBTITLE
A Survey of Terrain Modeling Technologies and Techniques

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT

6. AUTHOR(S)
J. David Lashlee, Nick Gorkavyi, and Jerry Snyder

5d. PROJECT NUMBER
ILIR

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Engineer Research and Development Center (ERDC)
Topographical Engineering Center (TEC)
7701 Telegraph Road,
Alexandria, VA 22135-3864

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/TEC TR-08-2

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Corps of Engineers
441 G Street NW
Washington, DC 20314-1000

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Test planning, rehearsal, and distributed test events for Future Combat Systems (FCS) require rapid generation of high-fidelity synthetic environments. These environments consist of high resolution synthetic scenes of test and training ranges, which use high and low resolution digital terrain surface models, 2-D and 3-D surface objects and other geospatial data to replicate site conditions. The largest component of developing synthetic 3-D scenes is the commercially available Interferometric Synthetic Aperture Radar (IFSAR) and LIght Detection And Ranging (LIDAR) data collected from airborne platforms. These industries are seeing rapid growth in data availability, and numbers and types of sensors, but the commercially available software used to process these types of data to provide high resolution, high accuracy topographic products is limited in its ability to process and produce data quickly and accurately for FCS. This work assessed the fidelity and quality of the commercial digital surface model (DSM) and digital terrain model (DTM) from Intermap; developed algorithms based on automated feature extraction (AFE) for LIDAR data that can be applied to processing IFSAR data (to perform foliage/vegetation removal and building/structure filtering while maintaining accurate terrain profile), and assessed improvements in constructing DSM/DTM using the Computational Consulting Services (CCS)-developed methods.

15. SUBJECT TERMS

Survey	Terrain Modeling	Future Combat Systems (FCS)
digital surface model (DSM)	LIght Detection And Ranging (LIDAR)	Interferometric Synthetic Aperture Radar (IFSAR)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	SAR	156	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.1